# Tuning dynamical properties of nanoscale systems via atomic-level modifications: insights from all-atom Molecular Dynamics simulations


**Universidad Autónoma de Madrid**
**Departamento de Física Teórica**
**de la Materia Condensada**

Dissertation submitted to obtain the degree of *Doctor en Física*

**Author:** María Ortega Cruz
**Supervisors:** Prof. Rubén Pérez Pérez / Dr. J.Guilherme Vilhena Albuquerque d'Orey

*Madrid, November 2019*

*A mis padres,*

*Look out 'cause here I come, and I'm marching on to the beat I drum,*
*I'm not scared to be seen, I make no apologies, this is me.*
*"This is me", The Greatest Showman*

# Agradecimientos

Esta tesis no hubiera sido posible sin la ayuda y el apoyo que he recibido durante todos estos años de mis dos directores de tesis, Rubén Pérez y Guilherme Vilhena. A Rubén, quiero agradecerle la confianza que depositó en mi cuando me aceptó en su grupo para hacer esta tesis, y que se ha mantenido a lo largo de estos años. Ha sido un verdadero placer trabajar a su lado por la cantidad de cosas que aprendes hablando con él y por el buen ambiente que genera, ya sea en el trabajo como fuera del mismo. Y a Guilherme,...creo que con solo estas líneas no es posible expresar todo lo que le agradezco. Desde su paciencia al enseñarme simulaciones de dinámica molecular; pasando por su cercanía; y hasta por aguantar mis "enfados", cuando había algo con lo que no estaba de acuerdo (ya se sabe... la confianza da asco). A pesar de ello, siempre me ha tratado con cariño, se ha preocupado de que siga trabajando a gusto y ha intentado hacer lo que creía que era lo mejor para mí. Todo esto me ha ayudado a crecer como científica sientiéndome muy arropada y casi como en casa. Así que, por todo ello, muchas gracias a los dos.

Mis agradecimientos se hacen extensibles a todos los miembros del grupo de investigación *SPM-TH* con los que he coincidido a lo largo de estos años: Diego, Lucía, Michael, Carlos, Linda, Pablo, Milica, Krzysztof, Ruth, Jaime, Emiliano, Salvatore y Alberto. Desde el primer día, me acogieron como una más y he estado muy agusto con ellos, además de ayudarme cuando lo he necesitado. Entre ellos, me gustaría destacar a Diego y Emiliano, con los que en algún momento he coincidido en el mismo despacho; a Salvatore y a Alberto, con los que comparto la suerte de estudiar propiedades físicas de biomoléculas; y por supuesto a Linda y a Carlos. Con estos dos últimos he tenido la fortuna

de poder colaborar en un proyecto de investigación, lo que me ha permitido corroborar que ademas de excelentes personas (que ya lo sabía) son fantásticos investigadores. Con ellos me he sentido super cómoda y valorada trabajando, lo que es de agradecer. Además, ambos, y especialmente Carlos, me han apoyado y animado mucho durante toda mi tesis, y más durante la escritura de la misma. Gracias por todos los "¿Qué tal lo llevas?" porque, aunque parezca poco, ayuda mucho, y hace que os valore mucho más como personas de lo que ya os valoraba.

Además, quiero agradecer a todos los miembros, presentes y pasados, del departamento de Física Teórica de la Materia Condensada. La amabilidad y el buen ambiente al que todos ellos contribuyen han hecho de este departamento un sitio muy agradable para trabajar. Entre ellos, quiero destacar a todos los que han pasado durante mis años de doctorado por el "*caniche-despacho*" (o despacho 402.3). A Javi del Pino, por sus consejos y animos durante mi primer e inocente año de doctorado; a Sergio, porque sus expresiones y su manera de hablar siempre conseguían endulzar las conversaciones que teníamos en el despacho; a Víctor, por su paciencia, por su apoyo, por sus viernes *fúsion*, por su eterna sonrisa, y por muchas otras cosas que hicieron que le cogiera más cariño del que ya le había cogido durante la carrera; a Rubén, porque aunque le gustaba "pinchar" mucho, siempre estuvo ahí cuando hizo falta; A Raúl, por acudir siempre a mis llamadas de auxilio, por ser una excelente persona y el mejor compañero de despacho posible; y a Alberto y a Juanpe, los cuales han tenido la paciencia necesaria para conocerme y aguantarme durante mis meses de escritura (que han sido los peores). Además, me gustaría agradecer también el apoyo que he sentido por parte de casi todos los estudiantes de doctorado del departamento (algunos de los cuales son ya doctores): Carlos Gonzalez, Rocío, Álvaro, Javi Galego, Nerea, Jose..., gracias por escucharme y por compartir vuestras experiencias del doctorado (y lo que no es el doctorado) conmigo. Por último, no puedo olvidarme de Laura y Stefan, cuya eficiencia siempre ha ayudado a que el camino de la tesis resulte más fácil.

No puedo dejar de mencionar a todas las personas con las que he colaborado (científicamente hablando) a lo largo de la tesis. Gracias a Willy, Cristina y a Julio, por todo lo que aprendí con ellos sobre grafeno y por hacerme sentir que

mi opinión también importaba; a Alex y Mauricio, por recibir con buena cara nuestras simulaciones de cápsides víricas (y esperar con mucha paciencia las siguientes); a Ismael, por el entusiasmo con el que ha recibido siempre nuestros resultados y por todo lo que he aprendido colaborando con él; y a Juan Carlos, por intentar acercar los conceptos relacionados con el transporte electrónico en uniones moleculares a personas (como yo) que conocen menos esos conceptos. También quiero agradecer a Pedro Tarazona, Julio Gómez, Marisela Vélez, Ismael Díez, Montserrat Calleja, Francisco Cao y Cristina Flors su disposición a formar parte de mi tribunal de tesis.

Dejando atrás la parte mas profesional, me gustaría primero agradecer a todo los grupos de amigos/as que han estado/seguido a mi lado durante estos años de doctorado. Tanto los buenos momentos como las partes mas duras de esta etapa han sido siempre mejores al compartirlos con ellos/as. Entre los diferentes grupos, me gustaría destacar a los miembros de "Lah Canariah" (o mis amigos de Materiales), como Esther, Denise, Fran, Jose, Dani, etc, por aceptarme tan bien y por continuar a día de hoy preguntandome qué tal me va; a mis amigas "de Getafe" Lorena y Tamara, por escucharme y estar siempre ahí; a los "Machangos", siempre disponibles para fantasticos viajes a Cuenca (en tren, por supuesto) o relajantes días de piscina (con chichón incluido); y por último (y probablemente, más importante) a todos los miembros de "una beca no da para..." (y agregados): Alex, Elena, Aitor, Gus, Iris, Javi Cuerda, Juanfran, David (ejem), Laura R, Mariano, Laura S y Sergio. Con estos últimos, he compartido casi todas las comidas en la universidad durante estos años, he viajado, he cenado, he llorado, me he quejado (excepto en febrero, por supuesto) y, sobretodo, me he reido muchísimo. Han sido los mejores compañeros que he podido tener en este viaje de hacer la tesis y espero de verdad que no nos "desperdiguemos" (ejem, Elena), porque son probablemente la segunda mejor cosa que me llevo de estos años. Entre ellos, me gustaría destacar a Laura S, Sergio e Iris, porque formar parte de "la primera promoción del Grado de Física de la UAM" une mucho; a Elena, porque escribir una tesis en verano es mucho menos duro si lo haces con una amiga como ella; y a Laura R, porque desde que nos dimos cuenta que teníamos el mismo portatil en primero de carrera hemos

llevado vidas "paralelas", y no sabes lo que me alegro de ello. Sin tu apoyo Laura, este "manos arriba" hubiera sido mucho más dificil.

Y ya, para ir finalizando, quiero mencionar a las personas que, no solo me han apoyado durante estos años de doctorado, sino que lo llevan haciendo durante toda mi vida: mi familia. A mis padres Pedro y Nieves, porque no se que habría hecho sin su cariño y apoyo incondicional. Ellos, ya sea en forma de abrazo, en forma de tupper o de "¿Qué tal, como estás?" han estado siempre ahí cuando lo he necesitado. Al mejor regalo que ellos me han dado, mi hermana Diana, por sus abrazos, sus risas, sus ganas de sorprender y hacer felices a la gente que le rodea, y, en definitiva, por ser la mejor hermana del mundo. A mi familia palmera, es decir, mis tíos Víctor y Blanca y mis guapísimas primas Raquel y Cristina. Ya es una suerte tener familia en la isla de La Palma, pero si encima son ellos, muchísimo mas. Gracias por acogerme este último mes de Agosto, pues conseguisteis que los pocos dias libres que tenía los disfrutara como si fueran el doble. Y, por último, a toda mi familia palentina (los "Ortegas por el mundo"), porque el "calor familiar" que me habeis transmitido durante todos estos años es difícil de medir. De entre toda la gente que pertenece a este último grupo, me gustaría destacar a mis abuelos, por seguir enseñándonos la importancia de cuidarnos en familia; a mi prima Verónica, porque se acaba de embarcar en esto de hacer un doctorado en física y porque es la alegría personificada; a mi tía Cristi, por seguir pendiente de mi y alegrarse de mis logros; y a mi ahijado Andrés, porque su sonrisa picara es la mejor del mundo.

Para terminar, quiero agradecer a David por aguantar mis rayadas, por intentar siempre sacarme una carcajada, por enseñarme a no juzgar y a expresar lo que me pasa y, sobretodo, por crear un hogar tan bonito y seguro para mí durante estos años de doctorado. Espero que lo sigamos manteniendo, aunque ahora nos toque hacerlo a bastantes kilometros de distancia.

## Table of terms

**2D** two–dimensional.

**AFM** Atomic Force Microscopy.

**BSA** Bovine Serum Albumin.

**CNTs** Carbon Nanotubes.

**CSA** contact surface area.

**DNA** Deoxyribonucleic Acid.

**EC** electrochemical.

**Fab** antigen-binding fragments.

**Fc** fragment crystallizable region.

**GB** Generalized Born.

**h-bond** hydrogen bond.

**h-bonds** hydrogen bonds.

**HV** High Vacuum.

**IgG** Immunoglobulin G.

**MD** Molecular Dynamics.

**MVM** Minute Virus of Mice.

**NMR** Nuclear Magnetic Resonance.

**NPT** constant pressure.

**NVT** constant volume.

**OPLS** Optimized Potentials for Liquid Simulations.

**PBC** Periodic Boundary Conditions.

**PME** Particle Mesh Ewald.

$\mathbf{R}_g$ radius of gyration.

**REBO** reactive empirical bond order.

**RMSD** root-mean-square deviation.

**RMSF** root-mean-square fluctuation.

**SAM** self–assembled monolayer.

**SASA** solvent accessible surface area.

**SMD** Steered Molecular Dynamics.

**ssDNA** single-stranded DNA.

**STM** Scanning Tunneling Microscopy.

**TEC** Thermal Expansion Coefficient.

**UHV** Ultra–High Vacuum.

**vdW** van-der-Waals.

**water/ssDNA** water per single–stranded DNA.

# Contents

# Abstract

The development in the 1980s of the first scanning probe techniques, i.e Scanning tunneling microscopy (STM) and Atomic Force microscopy (AFM), enabled the manipulation and characterization of very different systems at the nanoscale, such as 2D materials and biomolecules. This advent, among others, promoted the transition of nanoscience to technological applications, such as electronic components, sensors and biomedical sensing devices, with novel unique properties. Still, to improve upon their integration and performance, a detailed understanding of their physical properties is called for. This endeavor gave rise to novel research fields aiming to tune the properties of these nano-systems *atom-by-atom*, in a bottom up approach. The dramatic effect that changes at single atom level may have in these systems is better realized through the giant-stiffening of one of the worlds stiffest materials, i.e. graphene. In particular, by removing only a 0.2% of the carbon atoms composing the graphene one may increase its Young modulus over a factor of 60% , an effect which has no parallel at the macroscale. This and other results help to illustrate the importance of atomic-level interactions/effects in the properties of nano-systems.

In this view, theoretical modeling has proven to be a fundamental tool to augment the resolution/understanding of the experiments down to the single atom. Among the different simulation approaches, coarse-grain models have been widely explored to understand the properties of nanoscale materials. However, these models lack the atomistic detail required to understand the effect of single atom modifications. To circumvent this limitation, in this thesis we harness over 30 years of development of "all-atom" classical Molecular Dynamics (MD) simulations in the biophysics community to now study nanoscale systems

and in particular how single atom changes may drastically alter their properties. Notably tackling these (large) nanoscale problems has only been possible early this decade with the advent of new computing architectures such as graphical cards (GPUs) allowing to study systems that are several nanometers large and most importantly for very long periods of time (microsecond long). Here we aim to inspect not only the dynamical changes occurring in a nano-system by introducing atomic modification on its (intrinsic) structure, but also when these modifications are introduced on its environment (e.g. solvent conditions). The latter is particularly relevant in biological systems, where the solvent is known to actively modulate their function as in the case of *biological waters*. On the overall we are able to tackle many different open questions present in the literature such as: (1) the importance of the atomistic detail of water structuring in mediating the protein-surface interaction; (2) how water may be used to actively control both the structure and mechanical properties of an assembly of single-stranded DNA (ssDNA) molecules; (3) how the dynamical properties of a nano-system may be tuned by changing a small percentage of the atoms ($<1\%$) composing it.

The thesis manuscript is organized as follows: The first chapter outlines the previous experimental works motivating the thesis as well as the current state–of–the–art in modeling the role of atomic *mutations* at the nanoscale. Subsequently, we describe the different simulation approaches here employed.

In the second chapter, we show the importance of the atomistic description of the solvent molecules when in mediating the protein-surface interaction. This study was motivated by previous theoretical works, which systematically reported protein unfolding upon its adsorption when using implicit solvation methods (where solvent is described as a continuous dielectric medium). The advantage of these implicit solvation methods is that they allow us to reduce by almost a factor of 10 the number of atoms used in the simulations. Notwithstanding, a careful assessment of the validity of this approximation (replacing the water molecules interaction by an effective continuous approach) lacked an experimental validation. To that end, we focused on the adsorption process of a protein, the Inmunoglobulin G (IgG), over graphene with both explicit and

implicit solvation methods. Our implicit solvent simulation results show that the IgG protein unfolds upon its adsorption to graphene, which contrasts with both explicit solvent simulations and previously reported experimental studies. To unravel the origin of this contrasting behavior, we perform a detailed energy decomposition analysis of both implicit and explicit solvent simulations. The latter reveals that protein-surface attraction is counterbalanced by the cost of breaking the solvation shells surrounding both the protein and surface. As a result, the adsorption proceeds smoothly as is activated by the entropic water thermal fluctuations. Contrastingly, implicit solvent simulations underestimate the cost of breaking the solvation shells (as they are inexistent in such effective models), prompting an instantaneous abrupt IgG adsorption over graphene. This is accompanied by a pronounced loss of solute-solvent interaction resulting in a large electrostatic unbalance inside the protein which ultimately results in its unfolding.

In the third chapter, we focus on the effect of modifying the solute-solvent interactions on the mechanical properties of a biosystem. More specifically, we use MD simulations to study the variation of the mechanical response of a ssDNA self-assembled monolayer with its hydration (ranging from 0 to 600 water molecules per ssDNA). Note that the mechanical response is here quantified by the curvature of the ssDNAs assembly, being the highest curvature associated with larger stress. Our simulations show that the curvature of the monolayer goes through three well differentiated stages with increasing hydration (H), in agreement with previous experimental results: (1) at low hydration levels, the curvature increases with the hydration; (2) for H>100 water/ssDNA the curvature decreases until (3) it saturates at H=400 water/ssDNA. By performing a detailed energetic and structural analysis, we describe the physical mechanisms in these three stages: (1) the ssDNAs become stiffer with the hydration but they continue approaching to each other, prompting a more energetic (a larger) upward deflection of the monolayer as the hydration increases; (2) the ssDNAs length increases with the hydration to enhance the number of water molecules between them, which hinders the increase of the curvature; (3) the new water molecules added to the system remain outside the monolayer, and thus the struc-

tural and mechanical properties of the monolayer barely change. We also inspect the effect of the grafting density of the ssDNA monolayer on its hydration-driven mechanical response variation. All in all, our results reveal a surprising fact that water molecules play an active role in controlling the structural and mechanical properties of these assemblies.

Finally, in the fourth and fifth chapters we demonstrate how the dynamical properties of nano-systems may be tuned via *atom-by-atom* modifications of its intrinsic structure. In particular we focused on two different systems: the Azurin protein (chapter four) and graphene monolayers (chapter five). In the former (chapter four), we show that single amino-acid mutations albeit not affecting the structure result in a pronounced stiffening of the protein, most notably near its active site (the $Cu^{2+}$ center). This work emerged from a collaboration with STM experiments showing that the electronic transport process through a single gold-Azurin-gold junction sharply changes by replacing one of the amino-acids of the Azurin with a cysteine. Interestingly, our unrestrained molecular dynamics simulations of the protein and three mutants in water show: (1) their structure is practically identical to the "wild-type" form; (2)they fluctuate less that the "wild-type" protein. The second result indicates that the Azurin structure becomes stiffer with the introduction of such mutations. Additionally, from the adsorption process of both "wild-type" and mutant Azurin variants over a gold surface, we propose that quenching the Azurin fluctuations via point-mutation significantly alters its adsorption process over gold and therefore the gold-protein-gold contact. This result paves the way to understand how point mutations may affect the conductance in these molecular junctions.

In the fifth chapter, we inspect how a low density of monovacancies on graphene sheets can strongly alter its thermal expansion coefficient (TEC). Contrary to most materials, graphene TEC is negative, i.e. graphene contracts when heated. This abnormal effect is caused by the thermal excitation of the graphene out–of–plane vibration modes. These same modes are also quintessential to understand other graphene mechanical properties. Interestingly these have been postulated to explain the significant increase of graphene Young modulus through the introduction of small amount of defects. Still, an atomistic under-

standing of how these properties (out–of–plane vibrations, defects and mechanical properties) are related was missing. To unveil this connection, we perform 40 ns-long simulations of the graphene dynamics using different monovacancy densities (from 0 to 6 $\times 10^{12}$ cm$^{-2}$). Our results show a reduction of graphene TEC absolute value with the introduction of monovacancies, in agreement with experimental results. A detailed analysis of the stress distribution along the graphene membrane evidences that monovacancies create strain fields in their surrounding areas, which affects the low energy out–of–plane graphene modes responsible for the negative TEC.

## Resumen

El desarrollo en la década de los 80 de las primeras técnicas de sondas de barrido, entre las que se encuentran la microscopía de efecto túnel (STM) y la microscopía de fuerzas atómicas (AFM), permitieron la manipulación y la caracterización de una gran variedad de sistemas en la nanoescala, tales como los materiales 2D y las biomoléculas. Éstas técnicas, entre otras, estimularon la transición de la nanociencia hacia aplicaciones tecnológicas tales como componentes electrónicos, sensores y dispositivos de detección biomédica, todos ellos con propiedades únicas. Aún así, para optimizar su integración y rendimiento, es necesario entender detalladamente sus propiedades físicas. Esta necesidad dio lugar a nuevos campos de investigación con el objetivo de ajustar las propiedades de estos nano-sistemas átomo a átomo, bajo un enfoque "bottom-up". La mejor forma de ejemplificar el notable efecto que cambios a nivel atómico pueden producir en estos sistemas es a partir del intenso endurecimiento observado en uno de los materiales más rígidos del mundo, el grafeno. En éste material, la eliminación de tan sólo un $0.2\,\%$ de sus átomos puede incrementar su módulo de Young en un factor por encima del $60\,\%$, un efecto sin precedentes en la macro-escala. Este y otros resultados ayudan a ilustrar la importancia de las interacciones/efectos a nivel atómico en las propiedades de sistemas en la nanoescala.

En este contexto, la modelización teórica ha demostrado ser una herramienta fundamental para mejorar el entendimiento de los experimentos a escala atómica. De entre los distintos enfoques de simulación, los método de "grano-grueso" han sido ampliamente utilizados para entender las propiedades de los materiales en la nanoescala. Sin embargo, estos modelos carecen de los detalles atomísti-

cos requeridos para entender los efectos provocados por modificaciones monoatómicas. Para superar esta limitación, en esta tesis aprovechamos los más de 30 años de desarrollo de simulaciones atomísticas de Dinámica Molecular (MD) para estudiar sistemas en la nanoescala y, en particular, como cambios monoatómicos pueden alterar notablemente sus propiedades. Abordar este problema ha sido posible gracias a la aparición reciente de nuevas arquitecturas computacionales, tales como las tarjetas gráficas (GPUs), que permiten el estudio de sistemas de varios nanómetros y, más importante, durante largos períodos de tiempo (del orden de microsegundos). En esta tesis estudiamos no sólo el efecto de introducir modificaciones atómicas en la estructura intrínseca de un nanosistema, sino también cuando estas modificaciones son introducidas en el medio que los contiene. Este último caso es particularmente relevante en sistemas biológicos, en los que el solvente modula activamente su función y estructura. Con este estudio, somos capaces de abordar muchas cuestiones abiertas de distinta índole presentes en la literatura, tales como: (1) la importancia de los detalles atómicos a la hora de describir el papel del agua como mediador de las interacciones entre proteínas y superficies; (2) como el agua puede ser utilizado para controlar de forma activa la estructura y las propiedades mecánicas de monocapas autoensambladas de hebras simples de ADN (ssDNA); (3) como las propiedades dinámicas de un nano-sistema pueden ser controladas a partir de la modificación de un pequeño porcentaje ($<1\,\%$) de los átomos que lo componen.

El manuscrito de esta tesis está organizado de la siguiente manera: El primer capítulo explica, en términos generales, el trabajo experimental previo que motiva la tesis, así como las técnicas de vanguardia utilizadas en la modelización teórica del rol de mutaciones atómicas en la nanoescala. A continuación, se describe el metodo de modelización teorica que se ha utilizado en esta tesis (simulaciones de MD).

En el segundo capítulo, se muestra la importancia de la descripción atómica del solvente cuando este actúa como mediador de la interacción entre proteinas y superficies. Este estudio está motivado por trabajos teóricos previos, los cuales describieron sistemáticamente la perdida de la estructura terciaria de la proteina durante su adsorción al usar métodos de solvatación implícitos (en los que el

solvente se describe como un medio dieléctrico continuo). La ventaja de estos métodos reside en que permiten reducir hasta un factor 10 el número de átomos utilizado durante las simulaciones. A pesar de ello, es necesario comprobar que esta aproximación es válida para entender el proceso de adsorción de proteinas comparando los resultados a los que da lugar con resultados experimentales. Con este fin, estudiamos el proceso de adsorción de una proteína, la Inmunoglobulina G (IgG), sobre grafeno tanto con métodos de solvatación explícitos como implícitos. Los resultados obtenidos de las simulación de solvente implícito muestran que la proteína IgG pierde su estructura al adsorberse en grafeno, lo que contrasta tanto con los resultados obtenidos usando solvente explícito como con estudios experimentales previos. Para descifrar el origen de esta diferencia, se han analizado detalladamente las diferentes contribuciones energéticas a la adsorción en ambos tipos de simulaciones, es decir, usando solvente implicito y explicito. El análisis de estas últimas revela que la atracción energética entre la superficie y la proteina esta compensada por el coste energético de romper las capas del solvatación que rodean tanto a la proteína como a la superficie. Como consecuencia, la adsorción de esta proteina es un proceso lento y que solo ocurre gracias a las fluctuaciones térmicas entrópicas del agua. Sin embargo, las simulaciones por solvente implícito subestiman el coste de ruptura de las capas de solvatación, dando lugar a la adsorción instantanea de la IgG en el grafeno. Este proceso viene a su vez acompañado de una notable pérdida de interacción electrostática solvente-soluto que produce un gran desequilibrio electrostático en el interior de la proteína, que finalmente constituye el origen de la pérdida de estructura terciaria observada usando este tipo de solvatacion.

En el tercer capítulo, nos centramos en el efecto de modificar las interacciones solvente-soluto en las propiedades mecánicas de un sistema biológico. En concreto, se han utilizado simulaciones de MD para estudiar la variación de la respuesta mecánica de una monocapa autoensamblada de hebras simples de ADN (ssDNA) al cambiar su hidratación (desde 0 hasta 600 moléculas de agua por cada ssDNA). Nótese que la respuesta mecánica se cuantifica a partir de la curvatura de la monocapa, estando la curvatura máxima asociada con el estrés mecánico máximo. Nuestras simulaciones muestran que la curvatura pasa

por tres etapas bien diferenciadas cuando su hidratación (H) aumenta, lo que concuerda con resultados experimentales previos: (1) para niveles bajos de hidratación, la curvatura aumenta con la misma; (2) para valores de H >100, la curvatura disminuye hasta que (3) la evolucion de la curvatura satura cuando H >400. A partir de un análisis detallado de la estructura y de las interacciones energéticas que dan lugar al movimiento de la monocapa, describimos los mecanismos físicos detrás de cada una de estas etapas: (1) las hebras de ssDNA se endurecen con la hidratación y se aproximan las unas a las otras, dando lugar a que la monocapa se deflecte hacia arriba con más intensidad (deflexión mas grande) a medida que la hidratación aumenta; (2) la longitud de las hebras de ssDNA aumenta con la hidratación para incrementar el número de moléculas de agua que puede entrar entre ssDNAs adjacentes, lo que dificulta el aumento de la curvatura; (3) las nuevas moléculas de agua añadidas al sistema permanecen fuera de la monocapa, por lo tanto, las propiedades estructurales y mecánicas de la monocapa apenas se ven afectadas. También inspeccionamos el efecto de la densidad de hebras de la monocapa en las tres etapas anteriormente descritas. En general, nuestros resultados muestran el sorprendente hecho de que las moléculas de agua juegan un rol activo en el control de las propiedades estructurales y mecánicas de estos agregados de ssDNAs.

Finalmente, en los capítulos cuarto y quinto demostramos cómo las propiedades dinámicas de nano-sistemas puede controlarse mediante modificaciones "átomo a átomo" introducidas en su estructura. Particularmente, nos centramos en dos sistemas diferentes: la proteína Azurina (véase capítulo 4) y una monocapa de grafeno (véase capítulo 5). En el primer caso (capítulo 4), mostramos que la mutación de un solo aminoácido en esta proteina da lugar a un considerable endurecimiento de la misma, más apreciable en zonas cercanas a su centro activo (el centro $Cu^{2+}$). Este trabajo surge de una colaboración experimental en la que experimentos de STM muestran que el proceso de transporte electrónico a través de una unión oro-Azurin-oro cambia abruptamente al reemplazar uno de los amino-ácidos de la Azurina por una cisteína. Nuestras simulaciones de la dinámica libre en solucion de tres mutantes de la Azurina mostraron: (1) que sus estructuras son prácticamente idénticas respecto a las de la forma "wild-type" de

la Azurina; (2) que sus fluctuaciones son menores que en el caso de las proteína "wild-type". Este segundo resultado indica que la estructura de la Azurina se endurece con la introducción de estas mutaciones. Además, la simulación del proceso de adsorción sobre oro tanto de la estructura "wild-type" de la Azurina como de sus mutantes sugiere que la redución de las fluctuaciones de la Azurina a partir de la creación de estas mutaciones altera significativamente su proceso de adsorción y, por tanto, al contacto Azurina-Oro. Este resultado abre una vía para entender como mutaciones de un solo aminoacido pueden afectar la conductancia de la Azurina en este tipo de uniones moleculares.

En el quinto capítulo, analizamos como la baja densidad de monovacantes en láminas de grafeno puede alterar considerablemente su coeficiente de expansión térmica (TEC). Al contrario que la mayoría de los materiales, el TEC del grafeno es negativo, es decir, este se contrae al calentarse. Este extraño efecto es causado por la excitación térmica de los modos vibracionales "fuera de plano" del grafeno. Estos mismos modos también influencian otras propiedades mecánicas del grafeno. Cabe destacar que estos modos han sido postulados como el origen del notable aumento del modulo de Young del grafeno observado a partir de la introducción de una pequeña cantidad de defectos. A pesar de ello, aún no estaba clara la relación a nivel atómico entre estas tres propiedades (vibraciones "fuera de plano", defectos y propiedades mecánicas). Para explicar esta conexión, hemos realizado simulaciones de la dinámica del grafeno para diferentes valores de la temperatura y varias densidades de monovacantes (desde 0 hasta $6{\cdot}10^{12}$ $cm^{-2}$). Los resultados muestran una reducción del valor absoluto del TEC del grafeno con la introducción de monovacantes, como también se observa experimentalmente. A partir de un análisis detallado de la distribución del estrés a lo largo de las diferentes membranas de grafeno (diferentes densidades de monovacantes), observamos que las monovacantes crean campos de tensión en sus proximidades, lo que afecta a los modos "fuera de plano" del grafeno y, por lo tanto, al TEC.

# 1 General Introduction

Nanoscience is the scientific field which focuses on the manipulation and characterization of matter at the nanometer scale, i.e one millionth of a millimeter [1]. Many different molecular aggregates belong to that size scale, ranging from nanotubes and nanoparticles, over organic chemistry macromolecules, to Deoxyribonucleic Acid (DNA) and proteins. Thus, nanoscience is a interdisciplinary branch of knowledge that combines aspects of different basic sciences, such as chemistry, physics, biology, electronic engineering, medicine and biomedical engineering.

In the last decades, the ability to directly manipulate and characterize matter at the nanoscale via experimental investigations has been widely developed [1]. This capability has allowed to discover that the properties of nanoscale systems are significantly different from the ones shown by their mesoscopic counterparts (composed by the same atoms) [1]. For example, gold nano-clusters are excellent catalysts while their bulk counterpart is inert [2]. Moreover, carbon nanotubes present a combination of the features of the two naturally occurring bulk forms of carbon: strength (diamond) and electrical conductivity (graphite) [1]. Such astonishing divergence between the mesoscopic and nanoscopic levels reveals that understanding the properties at the mesoscale is not sufficient to describe material behavior. This points out the need to understand the physics rules that govern these systems in the nanoscale to enable their integration in different nanotechnological applications, such as electronic devices and sensors. In this thesis, we aim to theoretically study the dynamical properties of different systems at the nanoscale, and to understand how they can be tuned by just changing a small percentage of its atoms. This analysis contributes to gain insight into

the influence of atomic-scale modifications in the properties of nano-systems, helping to tightly control their properties, and paving the way for their device implementation and the improvement of their performance.

In this chapter, we first discuss the meaningful role of atomic modifications on the dynamical properties of different nano-structures and why understanding this role with theoretical simulations is a major computational challenge (see sec. 1.1). We focus on the two types of atomic-level alteration which have been analysed in this thesis, i.e. conducted over the system solvation conditions (extrinsic, see sec. 1.1.1), or over the intrinsic structure of the nano-system (see sec. 1.1.2). Then, we present the main features of the simulation method used in this thesis for studying this role : "all-atom" Molecular Dynamics (MD) simulations (see sec. 1.2). At the end of the chapter, an overview of this thesis is given (sec. 1.3).

## 1.1 Tuning dynamical properties at the nanoscale with atomic-level modifications: modelling requirements and state of the art

We here refer with *dynamical properties* to the characteristics of a system which describe its movement and deformation upon the influence of diverse stimuli, such as temperature changes (thermal properties), application of forces (mechanical properties) or the interaction with other systems. These properties are then essential for describing and predicting the behavior of a system in a given set-up and environment. Thus, gaining further insight about the dynamic properties of a system/structure is needed for allowing its potential use as active element of a given application. In fact, controlling these properties by altering the external or internal conditions of the system could help not only to anticipate its possible malfunction, but also to improve and increase its technological applicability.

At the nanoscale, the dynamical properties can be tuned by changing just a small percentage of the system/structure atomic interactions. This has been previously demonstrated in diverse fields, from biology [3] to materials science [4],

which evidences the importance of the atomic-level interactions in the characteristics of nano-systems. Still, the physical mechanisms by which a simple atomic modification can affect the general properties of a nano-system are not yet well understood. In this regard, theoretical modelling/simulations can serve as a critical tool to describe the role of these atomic alterations. This is not always a simple task, as changes on both the atomic interactions and the dynamical properties of the nano-system of interest must be properly captured by the chosen model. The former of these two requisites involves characterizing the dynamics of all the atoms of the system (achieved with atomistic models), which poses a larger computational cost than if a group of atoms is considered as a single interacting particle (as in coarse-grain models). The latter requirement, generally implies modelling the behavior/movement of a system larger than 10 nm, which increases even more the computational cost. The advent of new computing architectures, such us graphical cards (GPUs) or parallel CPU processing, has made possible to meet these two requisites/problems, as they have enabled to greatly reduce the simulation time, and, in turn, the computational cost. Before explaining in detail the characteristics of the simulation method used in this thesis , i.e. "all-atom" MD simulations, in the remainder of this section we present different examples in which the atomic influence on the dynamical properties of a nano-system has been evidenced, emphasising the main requisites needed to properly understand that influence with theoretical simulations. We have classified these examples according to the position where the atomic modifications where conducted, i.e. extrinsic (sec. 1.1.1) or intrinsic (sec. 1.1.2) to the structure of the system understudy.

### 1.1.1 Effect of extrinsic modifications

We refer here with *extrinsic* modifications to the ones performed in the environment of the system understudy. This type of modification is specially relevant in biological systems, as their native environment is a liquid solvent, i.e. a high percentage of water. In fact, the solvent not only acts as a merely passive agent in the properties of these biosystems, but it actively modulates their function and structure [5, 6]. For example, water gives rise to hydrophobic interactions

which stabilize the structure of globular proteins. Furthermore, the formation of an intricate hydrogen network around some biomolecules is needed to preserve their structure. This is the case of the single-stranded DNA (ssDNA) molecule, which can only maintain its helical structure when surrounded by water. Additionally, water molecules screen the charge-charge interactions, which reduces the electrostatic forces between charged groups and thus softens the interaction between them. All these reasons evidence the important role of water on the structure and properties of biomolecules and motivate the understanding of its influence on the dynamical properties of these systems.

In this regard, theoretical modelling can provide great insight. However, the explicit introduction of all the water molecules when modelling that influence requires a huge computational cost, as it typically increases the size of the system by, at least, a factor of 10. This has encouraged the fructiferous development of a manifold of implicit-solvation methods [7–11], which are based on representing the solvent as a continuous dielectric medium. These methods have the advantage of reducing the computational cost, but they lose the atomistic description of the solvent, which could be important, for example, for describing the solvation shells of the biomolecules. For that reason, these methods can only be used with a prior validation consisting on comparing their results with the ones obtained experimentally for an analogous system. In this thesis, we test the suitability of these methods for understanding a given dynamical property of a protein, i.e. its interaction with a hydrophobic surface (see more details in sec. 1.3 and on chap. 2). Moreover, we have inspected the effect of varying the hydration conditions of a complex bio-system, such as a self–assembled monolayer (SAM) of ssDNAs, on its mechanical properties. For this last problem, the necessity of tightly controlling both the amount of water in the system (up to a few water molecules) and the mechanical response of the whole SAM can only be achieved by combining the use of atomistic MD simulations and GPUs, thus enabling us to model the response of a large SAM when the humidity of the system is slightly increased (see more details in sec. 1.3 and chap. 3).

### 1.1.2 Effect of intrinsic modifications

We refer here with *intrinsic* modifications to those which involve changes on some of the structural elements composing a system, such as the introduction/presence of defects on the crystal structure of materials. The effect of this type of modifications has been explored in many different nanoscientific areas, ranging from materials science to medicine, evidencing that they could substantial influence not only the dynamical properties of nano-systems, but also their electrical [12, 13], magnetic [14] and biological [3, 15] properties. Among the different examples present in the literature, in this tesis, we focus on two types of *intrinsic* modifications: the introduction of defects on the structure of carbon-based nanomaterials, and the creation of single amino-acid mutations on proteins/protein complexes. In the reminder of this subsection, we detail the insights previously gained about these two intrinsic modifications with both experiments and theoretical modelling, and the open questions existent still in the literature about their effect.

#### Influence of defects on the mechanical properties of carbon based materials

Carbon-based nanomaterials, such as Carbon Nanotubes (CNTs) or graphene, have demonstrated impressive properties since their discovery (1990-2005). They have a Young's modulus of about 1 TPa [16, 17] and their tensile stress ranges between 60 GPa and 130 GPa [16, 17], which turns them into ones of the world's stiffest materials. Furthermore, they have exhibited other exceptional properties, such us a high thermal conductivity (2000-6000 $W \cdot m^{-1} \cdot K^{-1}$) [16, 18]. These and others outstanding properties evidence the great potential of these materials as active element of a wide range of nanotechnological applications, ranging from nano-sensors to field effect transistors [16].

A great deal of research effort has been devoted in the last decade to improve/control the properties of these carbon-based systems and thus promoting their implementation into nanotechnology devices. In this context, the influence of defects on their mechanical properties has been widely explored [16, 17, 21–

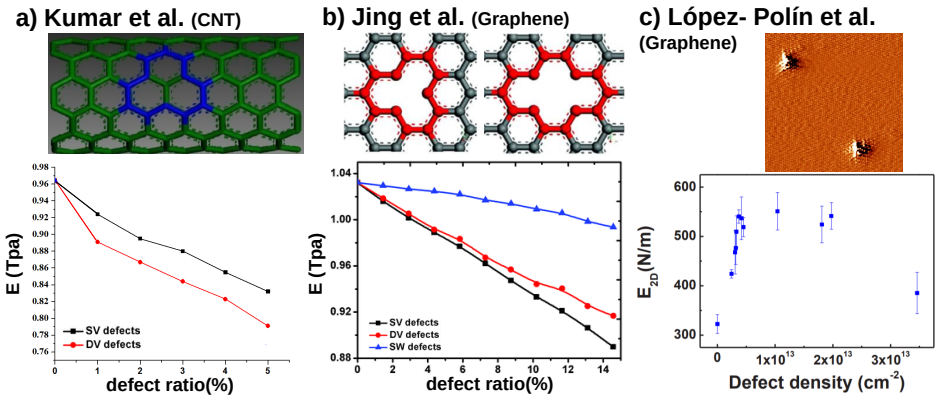



**Figure 1.1:** ***Influence of defects on the Young's modulus of carbon-based nanomaterials: results obtained from previous (a,b)MD and (c)experimental works.*** *From* ***a)*** *to* ***c)****, we displayed the Young's modulus variation obtained by increasing the number of defects in refs. 19, 20 and 4, respectively. The type of defects studied on each work is also here indicated, representing its form on the top panel. The results displayed in* ***a)*** *refer to the Young'modulus of CNTs, while the ones displayed in* ***b), c)*** *refer to graphene Young'modulus. These figures have been reproduced from ref. 4 by permission of Springer; from ref. 20 by permission of The Royal Society of Chemistry; and from ref. 19 by permission of Elsevier.*

23], as the presence of these intrinsic modifications is almost inevitable during the different production processes of these nanomaterials [16]. In fact, the first experimental works that analyzed the influence of defects on the mechanical properties of one of these carbon-based materials, i.e. graphene, were not especially designed for studying that influence, but the existence of defects was a direct consecuence of the graphene production process [22, 23]. In these experiments, they concluded that the Young's modulus of graphene decreases with the presence of defects on its structure [22, 23]. This result was confirmed by following theoretical works [19, 20, 24], where different kinds of defects, such as monovacancies, divacancies or Stone-Wales (SW) defects, were introduced on the crystal structure of these carbon-based materials, see Fig. 1.1a-b. As illustrated in this figure, they also predicted a monotonic reduction of the Young's modulus with defects density [19, 20], in agreement with one's intuition. However, in 2014, the first experimental work directly designed to introduce a

controlled density of monovacancies on graphene and evaluate how this change its mechanical properties concluded that its Young's modulus increases with a small density of monovacancies [4], see Fig. 1.1c. This striking result contrasts with previous experimental and theoretical results, which highlights the necessity of understanding the physical mechanisms underlying the effect of defects on graphene mechanical properties. Ref. 4 suggested that this counter-intuitive result may be arising from a defect-induced ironing of the "out-of-plane" thermal fluctuations of graphene, without further proof (see more details in sec. 5.1.2). To give further insight into this controversy, we here aim to inspect the effect of monovacancies on the "out-of-plane" thermal fluctuations of graphene by estimating another thermo–mechanical property of graphene influenced by these fluctuations, i.e. the Thermal Expansion Coefficient (TEC). For this estimation, we need not only a proper characterization of the structure of the monovacancy, but also to simulate the dynamics of sufficiently large graphene sheets for capturing changes on its "out-of-plane" fluctuations. This is then only possible with the combination of "all-atom" MD simulations (which provides the atomistic detail) and novel computing architectures, such as GPUs or parallel CPUs processing, which allows us to describe the dynamics of large graphene sheets during long-enough simulation times for determining changes on the graphene lattice parameter with very high precision ($10^{-5}$Å, see more details in chap. 5).

#### Influence of single amino-acid mutations on proteins

Proteins are large biomolecules conformed by one or more chains of amino-acid residues. An amino-acid is an organic compound composed by an amine group ($-NH_2$), a carboxyl group ($-COOH$) and a side chain. This last component is the one which changes from one amino-acid to another and its variety ranges from a solely hydrogen atom (alanine) to functional groups as complicated as an imidazol ring (histidine). The different amino-acids which conform a protein are sequentially bonded together via peptide bonds, i.e, a C-N covalent bond formed by the reaction between the carboxyl and amine grops of the two adjacent amino-acids, thereby forming the amino-acid chain (*primary* structure of the protein). This chain folds around itself, leading to the formation of different

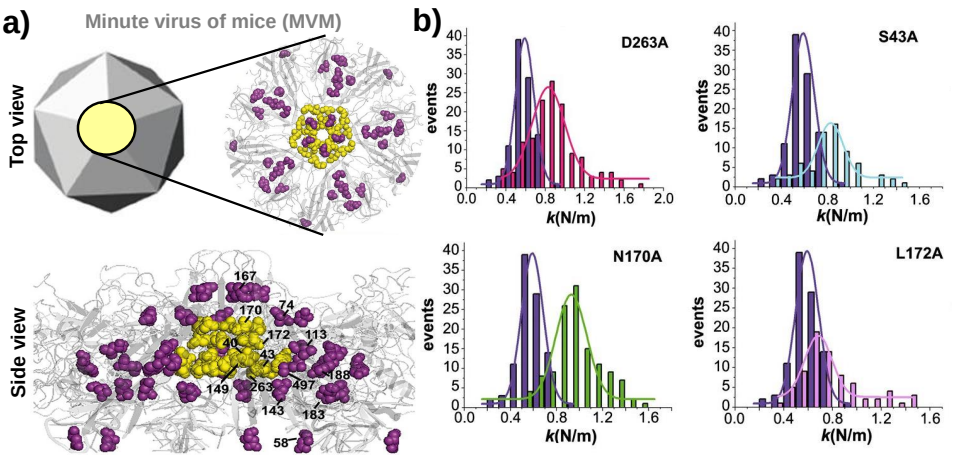



**Figure 1.2:** ***Analysis of the influence of single amino-acid mutations on the mechanical response of a virus capsid, the Minute Virus of Mice (MVM), performed in ref. 3 a)*** *Structure of the virus and position of the mutations. Here both top an side views of the capsid are shown.* ***b)*** *Distribution of the stiffness values (k) obtained through several Atomic Force Microscopy (AFM) indentations on each variant of the MVM capsid. In each graph, the distribution in purple corresponds to the wild-type capsid, and the one in a different color to a mutant capsid (labeled). These figures have been reproduced from ref. 3 by permission of the National Academy of Sciences.*

hydrogen bonds (h-bonds) between its different parts (*secondary* structure of the protein), forming patterns which ultimately determines the protein's unique 3D shape (*tertiary* structure of the protein). The secondary structures commonly known as $\alpha - helix$ and $\beta - sheet$ are examples of typical hydrogen bond (h-bond) patterns usually present in proteins.

The most interesting property of these biomolecules is its high specificity for performing successfully a given biological function, such as transporting the oxigen through the blood or separating the two strands of the DNA for the transcription of the genetic information. In fact, its structure is extremely linked to its biological function, as the loss of the former will imply that it could not perform the latter. This high specificity turns them into very interesting candidates to be incorporated as active elements of different biotecnological applications. However, proteins mostly need engineering at both its structure and

physical properties to meet the conditions required for their applications. This, together with the necessity of understanding the factors which could drive their malfunctioning, have encouraged the scientific community to study the properties of proteins and understand how they can be tuned by the introduction of modifications on its intrinsic structure. This endeavor has revealed that the dynamical properties of proteins can be drastically changed by the introduction of a single amino-acid mutation on its structure, i.e. by substituting one of its amino-acids by another [3, 15, 25–30]. In fact, not only their dynamical properties are influenced by this atomic-scale modification, but also their biological function [3, 15], which reinforces even more the necessity of understanding the effect of these subtle modifications. Additionally, this influence of single point mutations on the dynamical properties of proteins is also observed for more complex protein-based biosystems, such as virus capsids. Previous AFM experiments have demonstrated that the stiffness of the capsid of the MVM increases up to a 55% with the introduction of different single point mutations on each of the 60 sub-units that form the icosahedral capsid, see Fig. 1.2, without practically altering its structure [3]. These results have suggested that this might have some evolutionary origins, as stiffening the capsid structure may result in a reduction of its virological activity and, therefore, the softer capsides are naturally selected [31]. However, an understanding of how a single amino-acid mutation can affect the dynamical/mechanical properties of such large bio-systems is still lacking in the literature. We here give further insight into this problem by analyzing the influence of single amino-acid mutations on the dynamics of a small protein, the Azurin (see more details in sec. 1.3 and on chap. 4). Considering that large simulation times are required to accurately identify changes on the fluctuations of a protein ($>$ 100 ns), that influence can only be unveiled by combining graphical cards processing (GPUs) with the use of atomistic MD simulations.

## 1.2 Theoretical approach: atomistic MD simulations

To shed light on the influence of atomic modifications on the dynamical properties of the nano-systems described on the previous section, in this thesis we

have modelled their dynamics with "all-atom" MD simulations. This simulation method characterizes the dynamic of all the atoms of a system by numerically solving the Newton's equations of motion [32, 33]. In these simulations, the inter-atomic interactions are usually described by empirical or semi-empirical potentials, which are commonly known as *force fields* [32, 34]. In this section, we detail the main characteristics of the MD simulations performed in this thesis, such as the type of force fields used (see sec. 1.2.1) and other important modeling considerations (see sec. 1.2.2).

### 1.2.1 Force fields

The term *force field* refers to both the analytical expression which describes the potential energy of a system as a function of the coordinates of its particles and the set of empirical parameters entering into this expression [34]. These parameters are usually determined/estimated either from ab-initio quantum calculations or by fitting them to experimental data extracted from diverse techniques such as Nuclear Magnetic Resonance (NMR) spectroscopy, Raman spectroscopy or X-ray diffraction [34]. Ideally, they must be sufficiently accurate to reproduce the properties of interest and, at the same time, they must be simple enough not to increase excessively the computational cost. In parallel with the progress achieved in the comprehension of the physical mechanisms at the nanoscale with MD simulations, several force fields, with different degrees of complexity, have been developed and improved in the last decades to treat the inter-atomic interactions of many kinds of systems. The choice of a particular force field strongly depends on the system and the physical properties under study [32]. In this thesis, we have used two different force fields: the AMBER force fields and the bond-order potential AIREBO. The main characteristics of these force fields will be detailed in the following paragraphs.

#### Biomolecules: AMBER force fields

*AMBER* refers to a package of molecular simulation programs specialized on modelling and analyzing the dynamics of different biomolecules [35, 36]. Moreover, the name *AMBER* is used for referring to the inter-atomic empirical poten-

tials (force fields) that this package includes [36]. These force fields are primarily designed for the characterization of the dynamics of biomolecules, such as proteins and nucleic acids (*ff99SB* and *ff14SB* force fields), carbohydrates (*Glycam*) or other organic compounds (AMBER generalized force field) [36]. All of them have, in general, the following form [32, 35]:

$$\begin{aligned} E &= E_{bonded} + E_{non-bonded} \\ &= \sum_{bond} k_b(b-b_0)^2 + \sum_{angle} k_\theta(\theta-\theta_0)^2 + \sum_{dihedral} V_n \cdot (1+\cos(n\phi-\delta)) \\ &+ \sum_{LJ} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \sum_{eelec} \frac{q_i q_j}{r_{ij}} \end{aligned} \tag{1.1}$$

As shown in eq. 1.1, the AMBER force fields consider both *bonded interactions* (highlighted in magenta) and *non-bonded interactions* (highlighted in blue). The first term describes the vibrational characteristics of the covalent bonds present in the system and includes three different energetic interactions, see eq. 1.1: the *bond*, *angle* and *dihedral* terms. These three energetic interactions respectively describe the energy variation caused by changes in the bond length (stretching mode), the angle between two bonds (bending mode) or the dihedral angle between three bonds (torsional mode), see Fig. 1.3a. Concerning the *non-bonded interactions*, they characterize how the energy between two non-bonded atoms changes with their distance ($r_{ij}$), and are described via the sum of the Lennard-Jones potential (*LJ* term in eq. 1.1) and the electrostatic potential (*eelec* term in eq. 1.1). The latter describes the electrostatic interaction between charged atoms, referring $q_i$ and $q_j$ to the charge of the atoms involved on the interaction. The former (Lennard-Jones potential) includes both the Pauli repulsive interaction ($\propto r_{ij}^{-12}$) and the van-der-Waals (vdW) attractive interaction ($\propto -r_{ij}^{-6}$), which estimate the electrical quantum effects induced on two atoms for being very close to each other ($r_{ij} \sim 1-20$Å), see Fig. 1.3b.

The AMBER software, together with its force fields, have been widely used for characterizing important processes and properties of diverse biosystems, such as protein folding/unfolding [37, 38] or the mechanical response of nucleic-acids [39]. In fact, they have been cited more than 14000 times [36] since

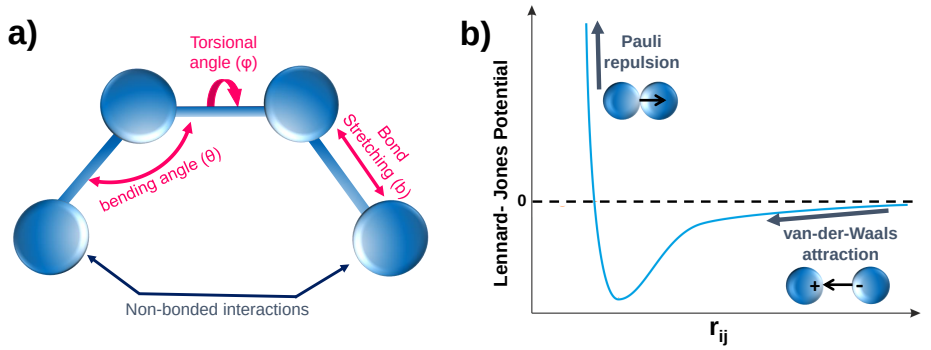



**Figure 1.3:** ***Characteristics of the empirical potential used in this thesis for modelling atomic interactions of biomolecules. a)****Schematic representation of the three vibrational effects of the bonds considered in this thesis via the bonded energy terms: stretching (bond term), bending (angle term) and torsion (dihedral term).* ***b)*** *Representation of the Lennard-Jones potential, highlighting the two physical interactions described by this potential (Pauli repulsion, vdW attraction).*

their emergence in the 2000s [35], which highlights their relevance in the context of biomolecular simulations. Furthermore, the software and, particularly, the force fields have been under continuous development during these last two decades [36, 40], thus establishing them as a modern and reliable way of theoretically describing the dynamics of biomolecules. For all these reasons, in this thesis we have used these state-of-the-art force fields and software for analysing the effect of atomic modifications on the dynamical properties of different biosystems (see sec. 1.3 for more details).

### Graphene out-of-plane vibration modes: AIREBO force field (LAMMPS package)

AIREBO is the second generation of a reactive empirical bond order (REBO) potential which successfully describes the intra-molecular interactions in carbon and hydrocarbon compounds [41]. This type of potentials (bond-order) have the advantage over other force fields that they are able to describe the different bonding states of an atom, i.e. they are able to distinguish for example between single and double bonds [41, 42] and describe the energetics involved in bond breaking. Thus, they can properly characterize chemical reactions involving the

breaking/remaking of a bond. There are many different bond-order potentials in the literature (Tersoff [43], Brenner [42], ReaxFF [44]) and all of them have the following form:

$$E_{ij}^{REBO} = E^{repulsive}(r_{ij}) + [b_{ij} \cdot E^{attractive}(r_{ij})] \quad (1.2)$$

where $E^{repulsive}$ and $E^{attractive}$ refer respectively to the repulsive and attractive pairwise interactions between two atoms ($i$ and $j$) separated a distance $r_{ij}$; while $b_{ij}$ represents the coupling of the $i-j$ bond with its chemical environment. This last term, i.e. $b_{ij}$, is the one which allows to include in the physical description of the bonds, the different chemical effects that may affect their strength, such as coordination numbers, bond angles or conjugation effects [41, 42]. The AIREBO potential extends the performance of one of this REBO potentials (Brenner [41]) by including a characterization of the torsional effects of the bonds (see Fig. 1.3 in sec. 1.2.1) and of the non-bonded interatomic interactions [41], i.e:

$$E^{AIREBO} = \frac{1}{2} \sum_{i} \sum_{j \neq i} \left[ E_{ij}^{REBO} + E_{ij}^{LJ} + \sum_{k \neq i,j} \sum_{l \neq i,j,k} E_{ijkl}^{tors} \right] \quad (1.3)$$

where the non-bonded interactions are described via the Lennard-Jones potential ($E_{ij}^{LJ} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}$) and the bond torsional effects ($E_{ijkl}^{tors}$) [41] are characterized analogously to the AMBER force fields (see eq. 1.1).

The AIREBO potential has demonstrated to accurately describe a wide range of mechanical [45] and thermal properties [46, 47] of pristine graphene [45–47]. In particular, the torsional term of this force field provides a good description of the low energy out–of–plane phonon modes of graphene (*ZA,ZO*), which are crucial to properly reproduce its thermal and mechanical response (see chap. 5 for more details). Furthermore, it has been also used for studying the mechanical properties of different carbon-based materials with defects [45, 47], demonstrating its validity for characterizing the influence of defects on these properties. For all these reasons, we have decided to use this potential for analyzing the effect of

atomic-scale modification on the thermo-mechanical response of graphene (see sec. 1.3 and chap. 5 for more details).In this case, we have used the AIREBO potential in combination with the Molecular dynamics package LAMMPS, which is more focused on materials modeling and includes this potential.

### 1.2.2 General features of our MD simulations

In this subsection, we provide a brief description of some of the state-of-the-art modeling techniques used in our simulations for improving their efficiency and for imposing some conditions to the dynamics of the systems under study. One of these methodologies is the application of Periodic Boundary Conditions (PBC). This technique consists on introducing all the atoms/molecules of a system into a 3D box with a convenient shape, e.g. a cubic box, and surrounding it with its translational images/replicas in the three spatial directions [32], see Fig. 1.4a. Thus, when an atom passes through one side of the *original* simulation box (unit cell), it enters with identical velocity into the opposite side of one of the images, which is equivalent to re-appearing at the other side of the unit cell (see Fig. 1.4a). This methodology allows to model the dynamics of a system genuinely conformed by a finite number of particles as it was infinite, which is specially relevant when simulating the dynamics of a system embedded on a liquid solution [32]. In this thesis, we use PBC on all the MD simulations where explicit solvent is used (see chap. 2 for more details) and for modeling the dynamics of a two–dimensional (2D) crystal such as graphene (see sec. 1.3 and chap. 5 for more details).

It is important to note that, using PBC, the spurious interaction between a given atom and its image could contribute to its dynamics. For avoiding this effect, we introduce a cut-off radius ($r_{cut}$, see Fig. 1.4a) to all the non-bonded interactions, i.e. the atom pairs which distances are greater than $r_{cut}$ have a null interaction. Note that the value of $r_{cut}$ has to be smaller than half the simulation box size [32] for avoiding the interaction with the images. With the proper choice of $r_{cut}$, we not only ignore this spurious interaction, but also we reduce the computational cost by considering only the non-bonded interactions that most affect the dynamics of a given atom (the strength of these interactions

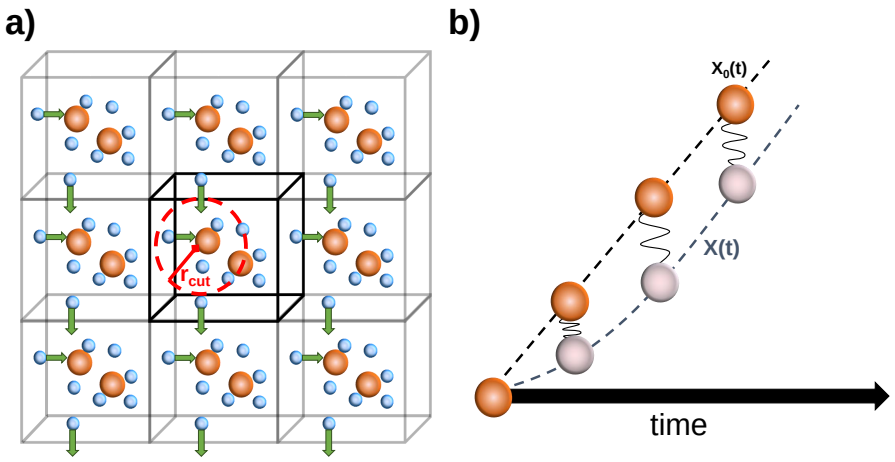



**Figure 1.4:** ***Schematic representation of the use of a)Periodic Boundary Conditions and b)Steered Molecular Dynamics.*** *In* ***a)*** *the molecules/atoms are represented via colored spheres, the* original *simulation box is represented in black, while the images are highlighted in grey. The mechanism whereby the atoms/molecules move from one side of the* original *simulation box to the other is here explained through green arrows. In* ***b)*** *we represent the difference between the trajectory of the* desired *value of the controlled generalized parameter ($X_0(t)$, black) with the trajectory of its* real *value ($X(t)$, blue). As shown in this figure, $X_0(t) = \frac{X_0(t_f)-X_0(t=0)}{t_f} \cdot t$*

decays with the distance, see eq. 1.1).

To control the temperature or/and the pressure of the system, we have used different thermostats and barostats. Regarding the thermostats, we have used the Langevin [48] and Nose-Hoover thermostats [49]. Both thermostats are based on introducing a fictitious variable to the force of each atom, whose physical meaning is similar to a friction, for slowing down or accelerating the atoms until the temperature of the system (estimated through the total kinetic energy) is equal to the desired value. When the temperature of the system is controlled and its volume is kept fixed the simulation is labeled as constant volume (NVT) simulation. When controlling the value of the pressure via a given barostat (e.g. Berendsen [50] or Montecarlo [51] , see more details in next chapters), the volume of the system changes in order to increase or decrease the pressure of the system (estimated through the total momentum interchanged by

the particles of the system) until the desired value. This type of simulation is usually known as constant pressure (NPT) simulation. In this thesis, we perform both NPT and NVT simulations (see next chapters for more details).

Finally, we explain the methodology followed when we need to restrain/control the dynamics of some parts of our system during the simulations. In this case, we have used Steered Molecular Dynamics (SMD) [52], a MD variant which is based on adding to the force field of the system an harmonic potential that influences only the trajectory of the group of atoms whose dynamics we want to control. This harmonic potential has the following form:

$$E = \frac{1}{2}K \cdot (X(t) - X_0(t))^2 \tag{1.4}$$

In eq. 1.4, $X_0(t)$, refers to the *desired* value of the generalized parameter of the system that is monitored (e.g. the distance between two atoms, the angle between three atoms, etc.) at each instant of the trajectory. Its time evolution is normally determined by fixing the simulation time ($t_f$) and the value of $X_0$ at the end of the simulation. The evolution is then defined considering that $X_0$ changes with a constant rate during the simulation, i.e. $X_0(t) = \frac{X_0(t_f) - X_0(t=0)}{t_f} \cdot t$, see Fig. 1.4b. The time evolution of the *real* value of that generalized parameter is defined in eq. 1.4 as $X(t)$ and strongly depends on the strength of the restrain, which is determined via the elastic constant $K$, see Fig. 1.4b. For small $K$, the restrain on the parameter $X$, will not be very effective and the difference between $X_0(t)$ and $X(t)$ will be larger. However, using a sufficiently large value of $K$, the value of $X$ can be tightly controlled at each time step of our simulations.

## 1.3 Thesis organization

In this thesis, we aim to understand with theoretical simulations the influence of atomic-level modification on the dynamical properties of different nano-systems. Once the importance and the state-of-the-art of this objective have been exposed, and the general characteristics of the chosen modelling approach ("all-atom" MD simulations) have been detailed, we proceed with the layout of the thesis.

The thesis is organized as follows:

- In chapter 2, we analyze the role of water in mediating biomolecule-surface interactions, focusing on understanding the relevance of an atomic description of the water molecules when these interactions are modeled.To that end, we study the adsorption process of a protein, Immunoglobulin G (IgG), over graphene with both explicit (all the water molecules are explicitly included) and implicit (the solvent is described as a continuous dielectric medium) solvation methods. Our implicit solvent simulation results show that the IgG protein unfolds upon its adsorption to graphene, contrasting with previously reported experimental results and also with explicit solvent simulations. Thus, the current description of these implicit solvation methods is not appropriate for correctly characterizing the protein adsorption process. By performing an energy decomposition analysis of both implicit and explicit solvent simulations, we unveil the two factors which cause the mischaracterization of this process with implicit solvation methods. Firstly, these methods underestimate the cost of breaking the surface and protein solvation shells; and, secondly, they omit the water-water electrostatic interaction gained by the solvent during adsorption. This study unveils the detailed energy balance between the different enthalphic interactions which contribute to the IgG adsorption and the relevance of entropic effects in this process.

- In chapter 3, we inspect the effect of the hydration conditions on the mechanical properties of a complex biosystem, i.e. a SAM of ssDNA. Our MD simulation results show that the curvature of this system changes more than a 25% with the hydration. More importantly, it goes through three well differentiated stages as the hydration increases, in agreement with previous experimental results: (1) the curvature increases with the hydration; (2) the curvature decreases with the hydration until (3) its variation saturates. To understand this significant hydration-drive mechanical effect, we perform a detailed energetic and structural analysis for all the hydration levels considered. Thus, we show that the gradual reduction of

the water-ssDNA affinity with the hydration causes significant changes on the ssDNAs structure, which in turn translate into a strong mechanical response variation of the whole SAM. Finally, we analyse how the grafting density value of the SAM influences its hydration-driven mechanical response variation. The results described in this chapter provide further insight about the effect of modifying the solute-solvent interactions on the mechanical properties of a DNA-based biosystem.

- Chapter 4 presents a study of the influence of single amino-acid mutations on the structure and dynamics of the protein Azurin. It has been experimentally shown that the electronic transport process through a metal-Azurin-metal junction sharply changes by replacing one of its amino-acids with a cysteine, which motivates us to understand the role of these *mutations* on its structure/dynamics. By simulating the unrestrained dynamics in water of three mutants studied in the experiments, we conclude that the introduction of these mutations barely changes the Azurin structure but it reduces its fluctuations, i.e. the Azurin becomes stiffer. Analyzing the adsorption process of both wild-type and mutant Azurin variants over a gold surface, we suggest that quenching the Azurin fluctuations with mutations may significantly alter its adsorption process over gold, even changing the Azurin most probable adsorption configurations. This result may explain the experiments, as variations on the Azurin-electrode configuration could ultimately affect the conductance through the junction.

- In chapter 5, we focus on the effect of atomistic modifications on the dynamical properties of a 2D material: a graphene monolayer. More precisely, we analyze how the introduction of a low density of monovacancies on graphene structure can strongly alter its large negative TEC. The negative TEC of graphene arises from the thermal excitation of the graphene out–of–plane vibration modes, which causes graphene to contract when heated. We here aim to understand how slight atomic modifications ($<0.2\%$) can influence these out-of-plane fluctuations by simulating the graphene dynamics using different monovacancy densities (from

0 to 6 $\times 10^{12}$ cm$^{-2}$) and estimating its TEC. Our results show that the graphene TEC absolute value reduces with the introduction of monovacancies, in agreement with experimental results. By inspecting the stress distribution along the different graphene sheets, we conclude that monovacancies create strain fields in their surrounding areas, which ultimately affect the low energy out–of–plane fluctuations of graphene.

- The conclusions of the work and the perspectives are presented in chapter 6.

# 2 Importance of the water model used to describe biomolecule-surface interactions

The surrounding solvent has an important role in mediating protein-surface interactions. Therefore, it is of paramount importance that the solvent methods employed to model this kind of processes are able to correctly capture the complex mechanisms occurring in the protein-water-surface interface. Recently, implicit solvent methods have been used in a regular basis to study different process where the water-protein-surface interaction plays a major role, such as protein adsorption. However, it is still not clear if these solvation methods appropriately describe the role of the water molecules during these processes [53]. In this chapter, we test the suitability of the two most popular implicit solvent methods based on the Generalized Born (GB) formalism, i.e. the HCT [54] and OBC [55] methods, to describe the adsorption process of the protein IgG on graphene with MD simulations. The results obtained using these two implicit solvent methods clearly show that the IgG denaturates upon adsorption, which contrasts with previously reported experimental and explicit solvent results [56]. To explain this discrepancy, we perform a detailed energy decomposition analysis of the adsorption process simulations using both explicit and implicit solvent methods. This analysis reveals that the ill-characterization of the IgG-graphene adsorption process using implicit methods arises from two main factors. Firstly, implicit solvation methods underestimate the cost of breaking both protein and substrate solvation shells, prompting an instantaneous adsorption process. Secondly, as the protein adsorbs, the implicit solvent description of the loss of solute-solvent interaction leads to large electrostatic energy unbalances inside

the protein, which causes its unfolding. These findings not only help to elucidate how implicit solvent models may be improved to accurately characterize the protein-surface interaction, but also reveal the detailed solvent mechanisms that control the adsorption of a globular protein such as the IgG to a hydrophobic surface.

## 2.1 Introduction

### 2.1.1 Protein adsorption

Protein adsorption has been a subject of great technological and fundamental interest [57, 58]. Proteins constitute the largest and most widely employed class of biomolecules for surface functionalization. As a result, numerous biotechnology applications such as biocompatible implants, biosensors and regenerative medicine [57, 59–61] rely on the process of protein adsorption. This broad applicability has stimulated in the last years the development of many experimental studies addressed to achieve the challenge of visualizing biomolecules and their hydration layers [62, 63] and controlling how a protein adsorbs to a surface [64–67]. However, the protein adsorption mechanism is not completely understood yet [58, 68, 69]. This results from the intrinsic nature of the process, that is driven by the interplay between both enthalpic and entropic forces [57]. The former are mostly composed by two components: the protein-surface interaction and the energy change due to protein's and water's structural rearrangement arising from its adsorption. The entropic forces, commonly known as hydrophobic forces, are, to a first approximation, a measure of how the disruption of the network of water-solute h-bonds induces or prevents the protein from adsorbing. Therefore, a proper description of hydrophobic forces requires a correct description of water ordering at the protein and surface interfaces[70, 71]. How each of these kinds of interactions (enthalpic/entropic) contributes in the adsorption process is still a very challenging problem, difficult to be determined solely from the experiments. For this reason, atomistic computer simulations, such as Classical MD, have turned into an essential tool to shed light on the protein adsorption process. After all, these simulations provide atomistic-level

insights into protein-substrate interactions thus giving a deeper understanding and control on protein adsorption experiments [72, 73].

### 2.1.2 Explicit and Implicit Solvation methods

In MD simulations of the dynamics of biomolecules, a correct description of the solvent environment is needed as it provides a crucial contribution to the structure and function of these systems [53], see sec. 1.1.1. The most direct way to properly account for the interaction between solvent-solute is by explicitly including all the water molecules present on the system. Nevertheless, this comes with a huge computational cost. By explicitly including all solvent molecules, one typically increases the system size by at least one order of magnitude, as compared to the size of the solute alone. Consequently, the exploration of the system phase space becomes slower, opening the possibility to an ill-characterization of the system final configurations simply because the equilibrium state has not been reached. This computational bottle-neck has encouraged a fructiferous development of a manifold of implicit-solvation methods [7–11]. In these methods, the solvent is represented by a continuous medium instead of individual explicit solvent molecules [53]. As a result, using this class of approximations one is able to greatly reduce the number of atoms of the solvated system. Therefore, the simulation computational cost is also reduced, allowing for a better exploration of the phase space while keeping a physically accurate description of the biological system of interest [73, 74]. The suitability and efficiency of these methods has long been validated on many different applications such as: binding free energies of small complexes to large proteins [75], protein folding landscapes [76, 77], or scoring protein conformations in fold prediction [78, 79]. Nevertheless, large differences between implicit models with experiments or explicit solvent approaches are generally expected when the properties of water molecules are substantially different from the bulk solvent, e.g. in media interfaces [77, 80]. The competitive interaction of water molecules for binding to the atoms of the molecule and the substrate surface is of paramount importance [68]. In addition, the formation of a quasi-ordered network of water molecules near both the substrate and the biomolecule surface [81] is misrepre-

sented with implicit solvation methods [80, 82–85]. As a result, these methods completely neglect energy terms that contribute to the molecule-surface binding such as the well-known dehydration process [57]. These and other known limitations [53] of implicit solvation methods seem to point out that the detailed role of solvent in mediating biomolecule-surface interactions can only be recovered by explicit solvation methods.

Although, implicit solvent methods have been used in a regular basis to study protein adsorption, it is still not clear if they are appropriate [53]. In fact, MD simulations of protein adsorption for the same or alike systems have shown qualitatively different results depending on the solvent method used. The works that use implicit solvent models [86–92] observe a severe degree of unfolding upon protein adsorption, which is in contrast with works were explicit solvent is used [56, 93–96]. This discrepancy was analyzed in a previous work [97], in which the adsorption dynamics of the Bovine Serum Albumin (BSA) on graphene using both implicit and explicit solvent models was studied. That work revealed a huge enhancement of the loss of the secondary structure of the adsorbed protein using implicit solvent methods. This result seemed to point out to a mischaracterization of the hydrophobic forces in implicit solvent methods. Usually this type of interaction is defined to be proportional to the solvent accessible surface area (SASA) of the system, which is a magnitude that changed drastically when the protein is near to the surface. Ref. [97] proposed that this drastic change could lead to an incorrect strong hydrophobic adsorption force, thus seriously enhancing the graphene-protein interaction and promoting a denaturation of the protein in the surface. However, that assumption was not correctly validated with a detailed analysis of the forces that participate in the protein adsorption process.

### 2.1.3 Overview

In this chapter, we test that assumption by comparing the performance of explicit and implicit solvation methods to simulate the adsorption process on a graphene surface of the IgG protein. The chapter is organized as follows. The rest of this section motivates the choice of IgG/Graphene as the system of study. In

sec. 2.2 the details about the MD simulations used in this chapter are described. Sec. 2.3 presents the final IgG-graphene adsorption configurations obtained with both implicit and explicit solvation methods, highlighting the similarities and differences with the experimental results reported in ref. 56. The assumption presented in ref. 97 is tested in sec. 2.4, analyzing in detail the role of the non-electrostatic solvation contribution when implicit solvent methods are used. Finally, in secs. 2.5 and 2.6, we perform a detailed analysis of the evolution of the vdW and electrostatic energy components of the system, respectively, for both explicit and implicit solvent simulations. This analysis highlights the importance of the entropic effects when a protein adsorbs on a hydrophobic surface, and shows the two factors that lead to the unfolding of the protein when implicit solvent is used.

### 2.1.4 IgG and its adsorption on Graphene

Antibodies, also named as immunoglobulins, are proteins which neutralize the antigens that enter into the human organism, such as viruses and bacteria [98]. Five different class of human antibodies are known, i.e IgA, IgD, IgE, IgG, and IgM, which differ in size, biological properties and function [98]. Among them, the IgG is the most common type of antibody found in the human body and is one of the most abundant plasma proteins. It is a *glycoprotein* composed by four peptide chains, i.e two heavy chains and two light chains connected by disulfide bonds, and two glycan chains, all of them arranged in a Y-shape [98], see Fig. 2.1a. Its molecular weight is near 150 kDa [98] and its secondary structure content is essentially constituted by $\beta - sheets$. This protein is characterized by two main regions, see Fig. 2.1a: a region whose structure remains practically unaltered when the IgG protein interacts with an antigen (*constant region*); and a region which is variable, thus allowing the protein to recognize and bind to each specific antigen via the antigen-binding sites (*variable region*) [98]. This high selectivity of the IgG towards antigens makes it an ideal candidate for numerous biomedical applications, such as immunoassays [99] and biosensors [100]. For these applications, the IgG must be previously immobilized on a substrate without losing its bioactivity and preferentially with the antigen-binding sites

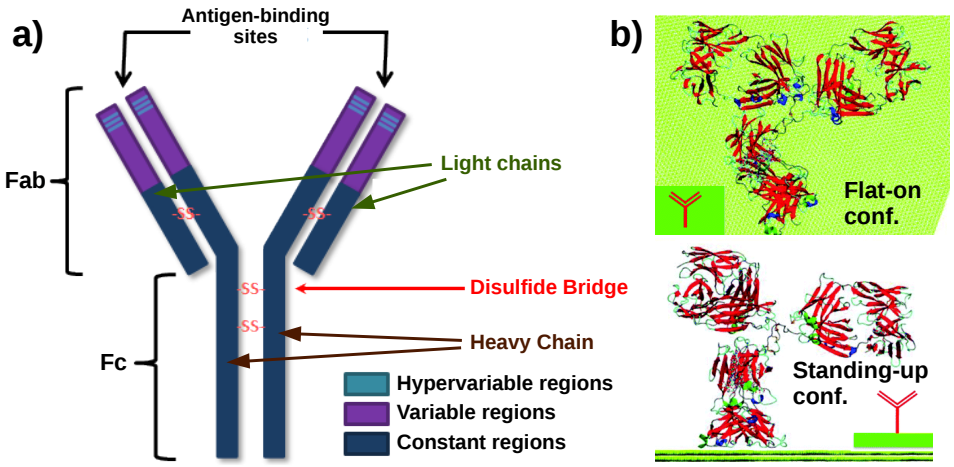



**Figure 2.1:** ***IgG structural characteristics.*** ***a)****Schematic representation of the IgG structure. The position of its four amino-acid chains and of the two regions (variable an constant) which characterize its structure are here depicted.* ***b)*** *Representation of the flat-on (top panel) and standing-up (bottom panel) configurations of the IgG on graphene.This figure has been reproduced from ref. 56 by permission of The Royal Society of Chemistry.*

pointing towards the solution. Since the IgG bioactivity is extremely sensitive to the conformational structure of their antigen-binding sites [57], studying and controlling its adsorption to surfaces is a matter of fundamental importance in the way of achieving IgG-based biomedical devices [56].

In last years, controlling the adsorption of the IgG has proven to be a difficult task as the IgG preferentially adsorbs flat on surfaces [56, 101–103], i.e. with the antigen-binding sites contacting the surface, see Fig. 2.1b. Consequently, these sites are less accessible when the protein is adsorbed and its bio-functionality reduces [56]. In a previous work [56], it has been validated via both AFM experiments and explicit solvent MD simulations that the IgG protein remains bioactive upon adsorption on graphene and that the most probable adsorption orientation of the IgG on graphene is a standing-up one, i.e. with its antigen-binding sites pointing towards the solution (see Fig. 2.1b). The preservation of its bioactivity indicates that both tertiary and secondary structure of the IgG are not significantly altered upon adsorption on graphene. These results unveil

the IgG-graphene system as a promising candidate for the next generation of biosensing devices [56].

## 2.2 Methods

### 2.2.1 System preparation and Force Fields

The protein structure of the IgG, composed of 1316 amino acids and 2 glycan heteropolymer chains, was obtained from the protein data bank (PDBID: 1IGT [104]). It is composed by four peptide chains (two heavy chains and two light chains, see Fig. 2.2a) arranged in a Y-shape. Protons were added to the protein structure according to the calculated ionization states [105] of its titrable groups at a blood pH of 7.4, resulting in a zero net charge. The IgG was then centered on top of a 20x20 $nm^2$ two-layer graphene slab with A-B stacking. We have oriented the IgG flat on the surface (see Fig. 2.2b-c). The initial distance between the lowest protein atom and the first graphene layer was 10 Å (see Fig. 2.2b) . All graphene atoms were set to be neutral and the bottom graphene layer of the slab was kept fixed during all simulations. This setup mimics the typical configuration in many adsorption experiments where a graphene layer is supported on an inert, mechanically rigid substrate.

The protein and the oligosaccharide were modeled by the AMBERff99SB [107] and Glycam04 [108] force fields respectively. These force-fields successfully sample the conformational space that an antibody explores in aqueous solution [109] and when it is adsorbed to surfaces [56]. The carbon atoms of the two-layered graphene were modeled by the Optimized Potentials for Liquid Simulations (OPLS) aromatic carbon force field present on AMBER's generalized force-field. [110]. This force field is known to properly describe graphene's mechanical and hydration properties [111], as well as its interaction with biological systems [109, 112]. Moreover, recent joint experimental and theoretical work, showed that this force field is capable of not only correctly characterizing the adsorption process of the IgG onto a graphene surface [56], but also to properly describe the graphene tribological properties in Ultra–High Vacuum (UHV) and in water conditions [113].

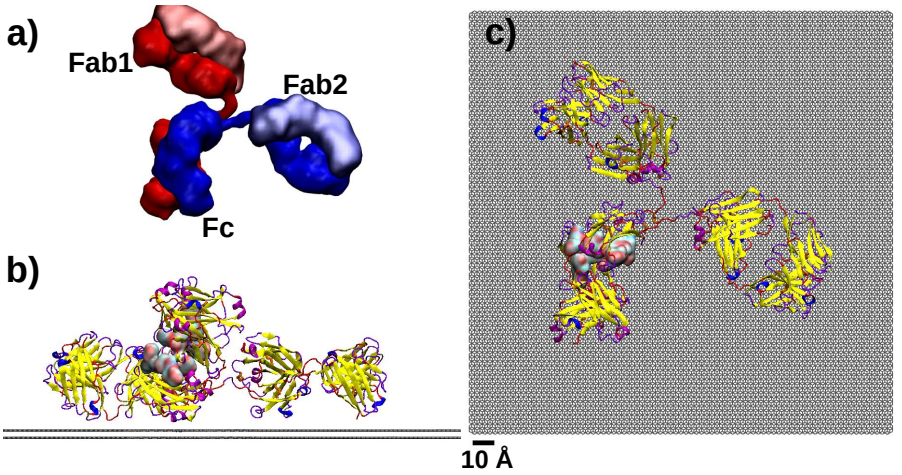



**Figure 2.2:** ***Initial configuration of the system IgG-graphene a)****Representation of the IgG molecule. Its four peptide chains have been marked with four different Connolly Surfaces [106]: the two light chains [104] are represented with metallic-pastel colors and the two heavy chains [104] with opaque colors. The three fragments that compose the IgG, i.e. the fragment crystallizable region (Fc) and the two antigen-binding fragments (Fab) are labeled accordingly.* ***b)****Side and* ***c)****Top views of the initial IgG-graphene configuration. The IgG is represented with its secondary structure: $\beta$-sheets (yellow), $\alpha$-helix (purple), 310-helix (dark-blue), turns (violet), and random-coils (red). The position of the two glycan chains has been highlighted with a Connolly Surface [106]. The disulfide bridges of the protein are also marked with an orange bond. The graphene has been represented in grey color.*

### 2.2.2 Solvation methods

We have treated the solvent via two different approaches: explicit and implicit solvation. In the explicit solvent case, the TIP3P water model [114] has been chosen. We have used PBC with a cubic unit cell that extends 20 Å beyond the molecule in the three directions. In order to solvate the system, this unit cell has been filled up with water molecules, placed in such a way that the minimum solute−water distance is 1 Å. This results in a system composed by ∼450.000 atoms. Note that in the implicit solvent simulations the system is only composed by the IgG and graphene atoms, i.e only ∼60.000 atoms.

In the implicit solvent case, two different methods based in the GB formalism

are used: the HCT [54] and OBC [55] methods. In this formalism, the total solvation free energy of a molecule is given by [55]:

$$E_{SOL} = E_{SOL(el)} + E_{SOL(nonel)}, \tag{2.1}$$

where $E_{SOL(el)}$ accounts for the variation in the electrostatic interaction between the molecule atoms due to the presence of a continuum dielectric solvent, and $E_{SOL(nonel)}$ includes both the solute-solvent vdW interactions and the hydrophobic effects. This second term is defined to be proportional to the total SASA of the molecule, i. e. $E_{SOL(nonel)} = \gamma SASA$ [54, 55, 115, 116]. As shown later in this work, (see sec. 2.4) the contribution of this term to the total solvent energy is negligible. Considering that, to reduce the simulation cost [117] and facilitate the interpretation of the results, we have not included this term in our implicit solvent simulations.

The first term, $E_{SOL(el)}$, is computed using the GB approximation [55], which describes the screening of the electrostatic interaction of two atoms of the molecule ($q_i$, $q_j$) due to the presence of a pure solvent with high dielectric value $\epsilon_\omega$ (80 for water at 300K) as:

$$\begin{aligned} E_{SOL(el)} &= E_{GB} \\ &= -(\frac{1}{2} - \frac{1}{2\epsilon_\omega}) \sum_{ij} \frac{q_i q_j}{r_{ij} \cdot \sqrt{1 + \frac{R_i^{born} R_j^{born}}{r_{ij}^2} e^{-r_{ij}^2 / 4 R_i^{born} R_j^{born}}}}, \end{aligned} \tag{2.2}$$

where $r_{ij}$ is the distance between atoms i and j, and $R_i^{born}$ is the effective Born radii of atom i. This radii reflects the degree of burial of the atom i inside the molecule, i.e. the deeper is the atom inside the molecule the higher the radii is. Therefore, according to eq 2.2, the atoms that are nearer to the molecular surface has a bigger contribution to $E_{SOL(el)}$ than the ones that are buried inside the molecule.

The difference between the two implicit solvent models here used is in the effective Born radii definition. In the HCT model [54], for computing the Born radii, the molecules that conform the solute are defined as a group of vdW spheres with dielectric constant unity. The rest of space, i.e. all the space that is not occupied by a sphere, has the high dielectric value of the media. The

Born radii for each atom is then computed as:

$$(R_k^{born})^{-1} = \rho_k^{-1} - \int_{\rho_k}^{\infty} \frac{dr}{r^2} H_k(r, \{\rho_{k'}, r_{kk'}\}_{all\ k'}) \tag{2.3}$$

where $H_k$ is the fraction of eclipsed surface area of a sphere of radius r centered at atom k with vdW radius $\rho_k$ when it is surrounded by spheres of radius $\rho_{k'}$ at a distance $r_{kk'}$. In the OBC [55] model the authors changed this original Born radii definition because it overestimates the solvation energy for deeply buried atoms. This overestimation is caused by the presence of artificial highly dielectric crevices in internal regions of the solute in which actually the solvent has been completely expelled [115].

### 2.2.3 MD parameters and Simulation protocol

We used the AMBER software suite [35] with NVIDIA GPU acceleration [118, 119]. For explicit solvent simulations, Particle Mesh Ewald (PME) [120, 121] (with a real–space cutoff of 10 Å) was used to account for long–range electrostatic interactions. vdW interactions were truncated at the real–space cutoff. SHAKE [122] algorithm was used to constrain bonds containing hydrogen atoms, thus allowing us to use an integration time step of 2 fs. Coordinates were saved every 1000 steps. A constant temperature of 300 K was ensured in all the simulations by means of a Langevin thermostat [48] with a friction coefficient $\gamma = 1\ ps^{-1}$ . For implicit solvent simulations, we do not use PBC, i.e. the long-range energetic contributions are not truncated. We also use the SHAKE [122] algorithm and keep a constant temperature using a Langevin thermostat [48] with the same conditions employed in explicit solvent simulations.

Our simulation protocol is composed by 3 main stages. In the first stage we performed an energy minimization to prevent steric clashes, using a combination of steepest descent and conjugate gradient methods. During this process, we kept weak restraints at the protein backbone and graphene substrate. In the second stage, we heated up the system from 0 to 300 K while restraining the position of the protein backbone and the first graphene layer. In the explicit solvent simulation, we use NPT conditions during this stage. A Berendsen barostat [50] with a relaxation time of $t_p = 1\ ps$ was used to keep the

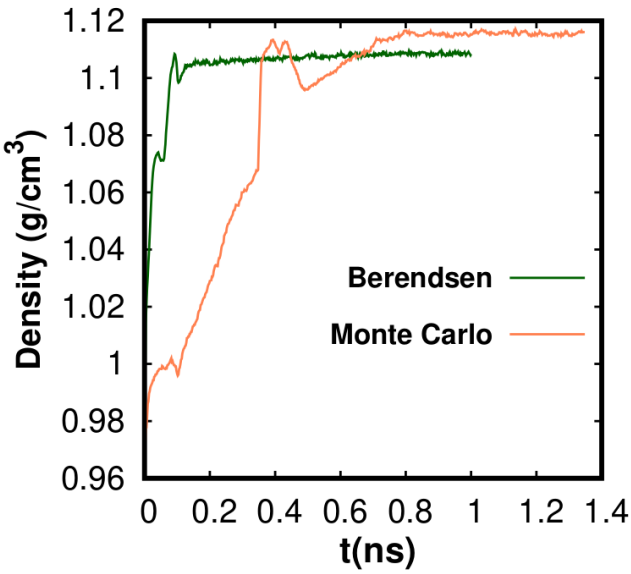



**Figure 2.3:** ***Evolution of the system density during the preliminary NPT simulation used for equilibrate the volume of the system with two different barostats.*** *During this preliminary stage, we heat up the system from 0 to 300 K while restraining the position of the protein backbone and the first graphene layer. Here we compare the results obtained with two different barostats: Berendsen [50] (the one used in our adsorption protocol, in green color) and Monte Carlo [51] (in orange color). As can be observed, the time evolution of the system density does depends on the barostat used, but the converged density value obtained at the end of the simulation ($\rho$) is roughly the same with the two barostast: $\rho_{Ber} = 1.1084\ g/cm^3$ and $\rho_{MC} = 1.1161\ g/cm^3$. Therefore, the difference between these two final values is less than a $1\%$.*

pressure constant at 1 atm. In spite of the known limitations of this barostat, ref. 123 shows that it can be efficiently used to equilibrate the system density. We have explicitly tested this point by analyzing the density evolution in that NPT preliminary stage using the Monte Carlo barostat [51]. Fig. 2.3 shows that the difference between the final density values obtained with these two methods is less than $1\%$, what reinforces the validity of the Berendsen barostat for equilibrating the system density.

Once these two preliminary stages had been performed, we started with the

adsorption process of the protein, in which we let the protein to freely adsorb to the substrate. This stage can be divided in three parts. First, we performed a 10 ns simulation in which the protein is free to adsorb to the nearby substrate. In the explicit solvent simulation, we use NVT periodic conditions during this first phase. Next, we performed a SMD simulation to enhance the adsorption process. This SMD process consisted in moving towards the surface a selected group of atoms at a constant velocity of 5 Å/ns via an harmonic restrain (with k=50 kcal/mol). This restrain is only applied to the alpha carbons belonging to 16 cysteine residues evenly distributed over the protein. Finally, once the protein has reached the surface, a long MD stage of up to 140 ns in a NVT ensemble was carried out for the explicit solvent simulation. In the case of implicit solvent simulations, the early high degree of unfolding justified to stop this final stage of our simulations at 30 ns.

### 2.2.4 Structural and energetic magnitudes analyzed

In order to characterize the protein-surface adsorption process, we have evaluated the following structural quantities: root-mean-square deviation (RMSD) of the protein backbone and domains [124], the inter-domains distances $d_{Fc-Fab1}$ and $d_{Fab1-Fab2}$ [124], the evolution of the secondary structure of the protein [125, 126] and the evolution of the contact surface area (CSA) between protein and substrate. The latter is calculated using the following definition:

$$CSA(t) = \frac{1}{2}\left[SASA_P(t) + SASA_S(t) - SASA_{P-S}(t)\right] \ , \qquad (2.4)$$

where the time dependent solvent-accessible-surface-area $SASA(t)$ [106] was calculated for protein ($SASA_P(t)$), graphene substrate ($SASA_S(t)$), and protein-substrate ($SASA_{P-S}(t)$).

We have also performed a detailed energy decomposition analysis to understand the differences between explicit and implicit solvent results. The only enthalpic interaction between protein and substrate is the vdW interaction, as the graphene atoms are not charged. When we refer here to *vdW interaction*, we mean the interaction described by the Lennard-Jones potential, which includes both the vdW attractive effects and the Pauli repulsive effects (see sec. 1.2.1).

We computed this interaction energy as a function of time via an energy decomposition on each frame of our MD trajectories. As a result, protein-surface interaction is given by:

$$E_{total}^{P-S}(t) = E_{vdW}^{P-S}(t) = E_{vdW}^{P+S}(t) - E_{vdW}^{P}(t) - E_{vdW}^{S}(t), \tag{2.5}$$

where $E_{vdW}^{P+S}(t)$ stands for the vdW energy of the combined protein+substrate system, and $E_{vdW}^{P}(t)$ and $E_{vdW}^{S}(t)$ are the vdW energies of the isolated protein and substrate, respectively. Moreover, to elucidate the role of the solvent in the adsorption process when explicit solvent is used, we have calculated also the interaction of each individual system i.e. the protein and the substrate, with the water molecules. These two magnitudes are computed analogously to $E_{vdW}^{P-S}$:

$$E_{total}^{W-P}(t) = E_{total}^{W+P}(t) - E_{total}^{P}(t) - E_{total}^{W}(t); \tag{2.6}$$

$$E_{total}^{W-S}(t) = E_{total}^{W+S}(t) - E_{total}^{S}(t) - E_{total}^{W}(t) \tag{2.7}$$

where $E_{total}^{W+P(S)}(t)$ stands for the total energy of the combined water+protein(surface) system, and $E_{total}^{P(S)}(t)$ and $E_{total}^{W}(t)$ are the total internal energies of the isolated protein(substrate) and water systems, respectively. The water-protein interaction is a combination of electrostatic and vdW effects, i.e. $E_{total}^{W-P}(t) = E_{el}^{W-P}(t) + E_{vdW}^{W-P}(t)$, while the water-substrate interaction is a pure vdW interaction as the graphene atoms are not charged, i.e. $E_{total}^{W-S}(t) = E_{vdW}^{W-S}(t)$. Finally, to reveal the solvent influence in the IgG-graphene adsorption process when implicit solvent methods are used, we have analyzed the evolution of $E_{SOL(el)}$ and $E_{SOL(nonel)}$, i.e. the electrostatic and non-electrostatic components of the system solvation energy (see sec. 2.2.2).

## 2.3 Structural dynamics of the adsorption: implicit vs explicit solvation models

### 2.3.1 IgG tertiary structure

Explicit solvent simulations results, see Fig. 2.4a, show that, upon adsorption, the IgG keeps its three-lobe structure in which each lobe corresponds to a fragment (Fab1/2, Fc). This structure is also clearly distinguished in previously

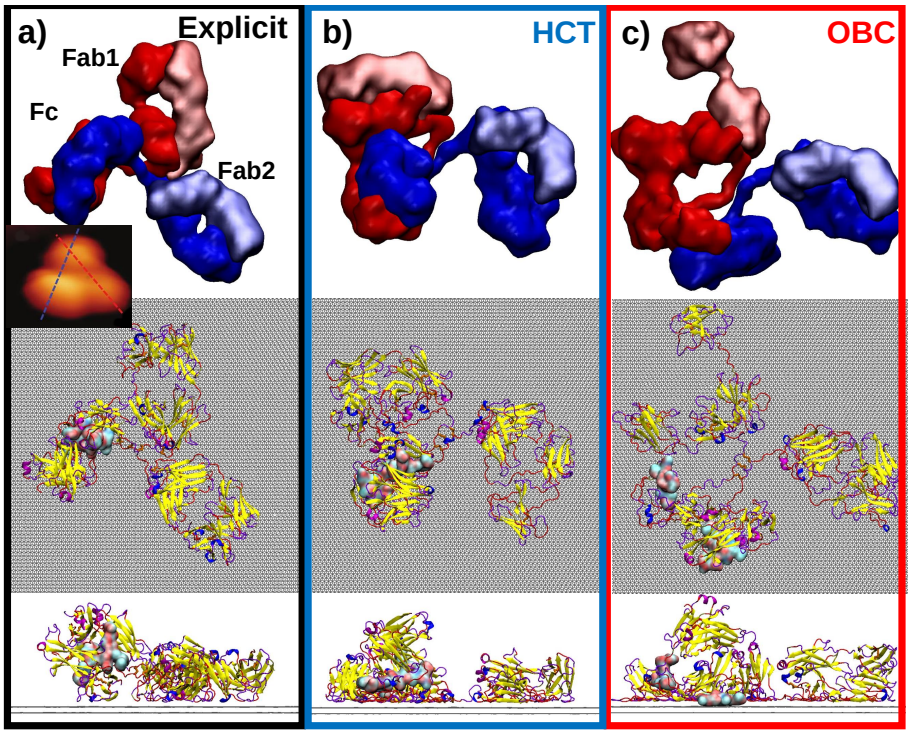



**Figure 2.4:** ***Final IgG-graphene configurations obtained at the end of the adsorption process using a)Explicit TIP3P [114], b)HCT implicit [54] and c)OBC implicit [55] solvation methods.*** *Both side (bottom row) and top views (middle row) of these configurations are shown. The color representation used for the system protein-substrate is the same as in the Fig. 2.2. The relative configuration of the four chains [104] that conforms the IgG upon adsorption is also showed (top row, see Fig. 2.2 for more details). The inset in* ***a)*** *shows an AFM image of the IgG antibody adsorbed in graphene published in ref. 56 and is reproduced by permission of The Royal Society of Chemistry.*

reported AFM experiments of IgG molecules adsorbed over graphene [56] (see inset in Fig. 2.4a). Furthermore, these experiments [56] show that one of the fragments is slightly higher than the other two. This is reflected by a brighter color of one of the lobes on the image shown in the inset of Fig. 2.4a. This observation is fully consistent with our MD explicit solvent results, as in our simulations the Fc fragment is slightly higher than the Fab fragments (see side views in Fig. 2.4a). Moreover, the inter-domain distances obtained in the AFM

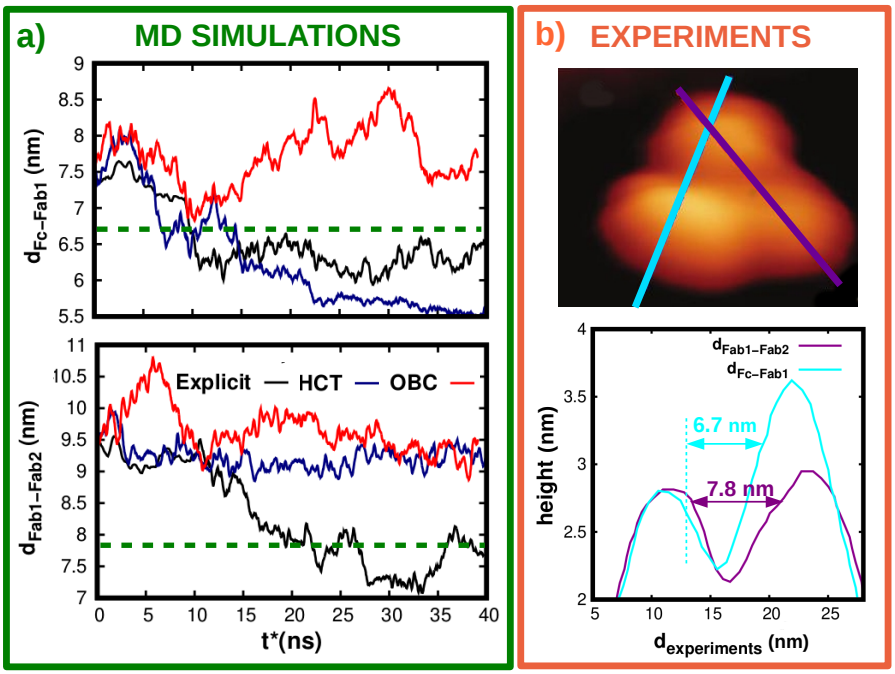



**Figure 2.5:** ***Time evolution of the inter-domain distances during the adsorption process. a)*** *Time evolution of the inter-domain distances $d_{Fc-Fab1}$ and $d_{Fab1-Fab2}$ using the three solvent methods here considered: Explicit [114] (black), HCT implicit [54] (blue) and OBC implicit [55] (red). The experimental values of these two inter-domain distances reported in ref. 56 are included in this figure with a green straight line.* ***b)*** *The method used for estimating the experimental inter-domain distances is here shown. This image is reproduced by permission of The Royal Society of Chemistry.*

experiments, i.e. $d_{Fab1-Fab2} \sim 7.8$ nm and $d_{Fab1-Fc} \sim 6.7$ nm, are also in good agreement with explicit solvent MD simulations [56] as shown in Fig. 2.5. Everything considered, we may conclude that the IgG adsorbed structure obtained in explicit solvent simulations is in excellent agreement with AFM experiments [56].

A key point concerning the IgG adsorption to surfaces is whether it preserves its biological functionality, i.e. the ability to bind specifically to an antigen via the Fab fragments.This is the case for IgG adsorbed to graphene as confirmed in ref. 56 via single molecule force spectroscopy experiments. This bioactivity can only occur if and only if the variable region of the Fab fragments preserves its tertiary structure [127]. Our explicit solvent simulation results corroborate

this, as the relative position of the two chains conforming each Fab fragment is preserved upon adsorption (see Fig. 2.4a).

In Fig. 2.4b-c, we report the final adsorption configurations obtained for both implicit solvent methods here considered, i.e. HCT [54] and OBC [55]. For both implicit solvent methods, we observe that the IgG tertiary structure alters significantly upon adsorption, which is in striking contrast with explicit solvent simulations results. Concerning the HCT [54] method results, see Fig. 2.4b, we identify three major structural changes. The first is the major unfolding of the Fab fragments, being the Fab2 the most affected. Secondly, the collapse of the Fc fragment occurring due to the clustering of the glycan chains located in the core of the Fc. The third and last aspect is the clustering of the Fab1 and Fc fragments observed in Fig. 2.4b, which is driven by the self-interaction of one of the heavy chains of the IgG. Considering now the OBC [55] results, see Fig. 2.4c, we note that the IgG also suffers a major unfolding due to adsorption, even more pronounced than the one obtained using the HCT method. The severe loss in the tertiary structure of Fc and Fab1 fragments leads to a spreading of the IgG over the surface. Additionally, the two glycan chains of the Fc fragment separate during the OBC implicit solvent simulation, see Fig. 2.4c

Both OBC and HCT structural changes are at odds with AFM experiments previously mentioned [56]. They are unable to reproduce the three lobe structure and the compact adsorption area of the IgG obtained in AFM experiments [56], see inset of Fig. 2.4a and Fig. 2.5. In addition, the major degree of unfolding of the Fab fragments observed in both implicit solvent simulations is incompatible with the bioactivity of the IgG observed experimentally [56]. This seems to indicate that OBC and HCT implicit methods do not correctly describe the adsorption process of the IgG.

All these changes in the IgG tertiary structure can be quantitatively confirmed analyzing the evolution of the RMSD deviation of each IgG fragment, which is shown in Fig. 2.6. Explicit solvent simulations show RMSD values below 5 Å for all fragments, highlighting the strong similarity of the final adsorbed configurations with its crystal structure [104]. This is not the case for implicit solvent results, which show for the three fragments larger RMSD values than

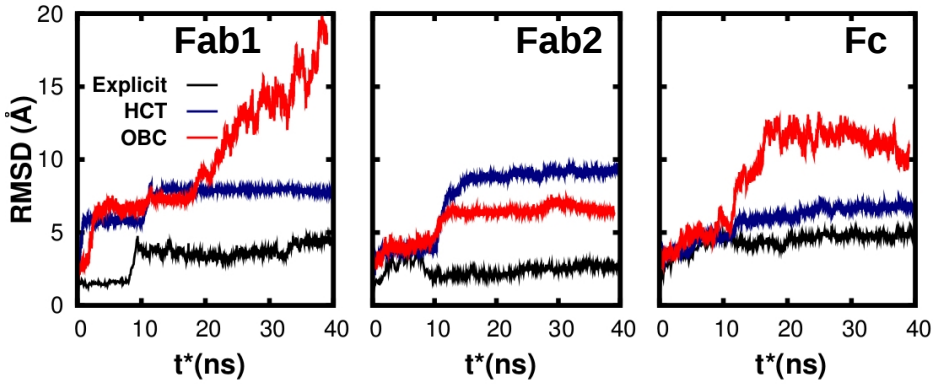



**Figure 2.6:** ***Changes in the IgG tertiary structure during the adsorption process.*** *Time evolution of the RMSD for the atoms belonging to the two antigen-binding fragments (Fab1 and Fab2 and the Fc fragment using the three different solvent methods here considered: Explicit [114] (black), HCT implicit [54] (blue) and OBC implicit [55] (red). The time of simulation for the explicit solvent case has been rescaled using the relation* $t* = t_{sim}/4.6$.

using explicit solvent. In fact, for both implicit solvent methods, the Fab1 fragment loses more tertiary structure (RMSD$\sim$ 7 Å) in the initial 10 ns (before the enhanced adsorption) than in 150 ns of the explicit solvent simulation. Similar processes may be observed in the other fragments of the molecule, evidencing the striking difference between implicit and explicit solvent methods in describing the IgG tertiary structure upon adsorption.

When comparing OBC [55] with HCT [54] MD results we observe that, on the overall, the former leads to a larger structural change. In particular, in Fig. 2.6, we see that the RMSD of Fab1 and Fc is not stabilized in 40 ns of OBC simulation, indicating that further unfolding would occur. Furthermore, the RMSD is above 10 Å for two fragments (Fc,Fab1), which is consistent with the almost complete loss of its tertiary structure shown in Fig. 2.4c. Considering that the sole difference between both implicit solvent methods is the electrostatic solvent energy contribution (see sec. 2.2.2), these differences seems to suggest that the observed protein unfolding upon adsorption may be related with the description of that energy component. In particular with the definition of the

degree of *burial* [55] of the aminoacids at the protein surface interface.

### 2.3.2 IgG secondary structure

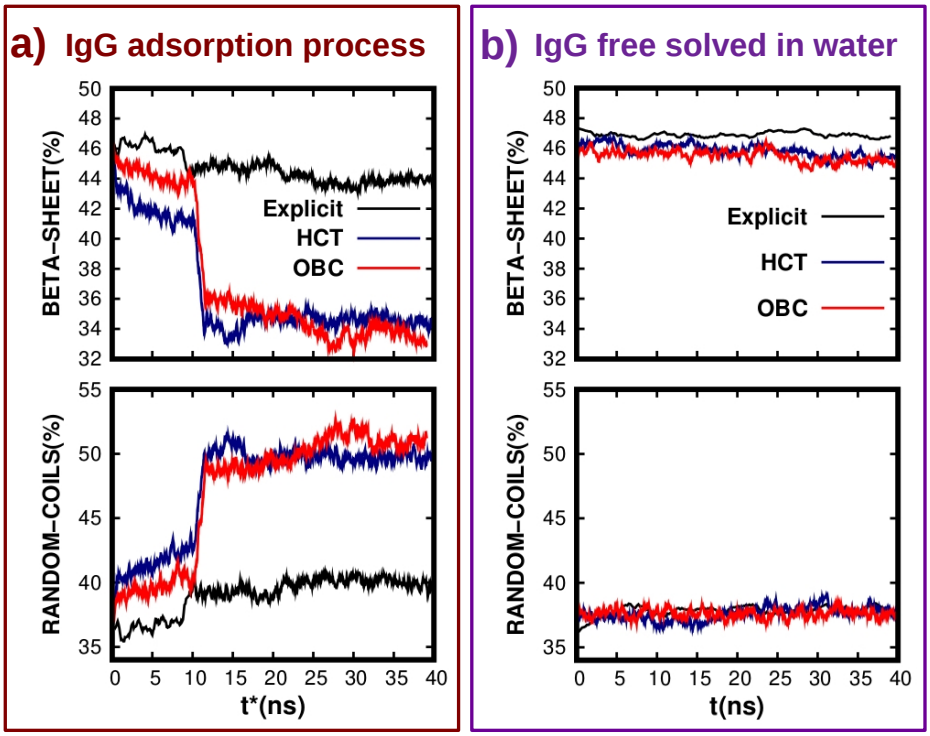



**Figure 2.7:** ***Time evolution of the IgG secondary structure content during the adsorption process. a)****We show the time evolution of the $\beta$-sheet (top) and random-coils (bottom) content for the three solvent methods here considered: Explicit [114] (black), HCT implicit [54] (blue) and OBC implicit [55] (red). The time of simulation for the explicit solvent case has been rescaled using the relation $t* = t_{sim}/4.6$.* ***b)****Time evolution of the $\beta$-sheet (top) and random-coils (bottom) content during a 40 ns long MD simulation of the IgG protein free solved in water. The results obtained with each one of the three solvent models here used, i.e. Explicit [114], HCT implicit [54] and OBC implicit [55], are shown. The color scheme is the same as in* ***a)****.*

To quantify the loss of protein secondary structure resulting from the adsorption, we analyze the evolution of the $\beta$-sheet content, which is the most abundant secondary structure in the IgG. The explicit solvent simulation pre-

dicts a decrease of the $\beta$-sheet content bellow 2 $\%$, see Fig. 2.7a. That result is consistent with the preservation of the functionality of the protein [127] reported experimentally [56]. However, implicit solvent simulations predict a $\beta$-sheet content decay higher than 10 $\%$, balanced with an increase of the content of random-coils as shown in Fig. 2.7a. These differences in the IgG secondary structure evolution do not happen when we simulate its free dynamic in solution using OBC and HCT implicit solvation methods, see Fig. 2.7b. Thus, the loss of the IgG secondary structure content observed in Fig. 2.7a can only arise from a deficient description of the adsorption process when those implicit solvation methods are used.

It is worth to note that the loss of $\beta$-sheet content observed in implicit solvent simulation results, see Fig. 2.7a, is not a gradual process. In the first 10 ns of implicit solvent MD simulations, the protein only lose a 3 $\%$ of its $\beta$-sheet content. However, when we enhance the adsorption using a short 1.6 ns long SMD simulation, see sec. 2.2.2, $\beta$-sheet content decreases more than a 7 $\%$. That abrupt change can be understood as the direct consequence of the great boost of the protein-surface contact induced in this SMD process. Still, the influence of the structural restrain imposed during this process, i.e. the distance of 16 cysteine residues, must be thoroughly analyzed to assure that it is not the origin of the sharp $\beta$-sheet loss observed in this stage. To that end, we perform the same SMD simulation but starting from a larger protein-surface distance ($d^{ini}_{P-S} = 30$ Å, see sec. 2.2.1). Thus, the protein is farther away from the surface and the protein-surface contact can not be the origin of an abrupt $\beta$-sheet content decrease. For performing this simulation, we use the OBC implicit solvent method [55], as that is the solvent method which distorts more the protein secondary structure during the adsorption process (see Fig. 2.7a). The results of that simulation are shown in Fig. 2.8. As clearly shown in this figure, when the protein does not interact with the substrate, the secondary structure of the IgG barely changes during the SMD process. This indicates that the restrain applied to 16 cysteine residues in this SMD stage is not affecting the IgG structural stability and then the abrupt secondary structure loss observed in Fig. 2.7a is only promoted by the increase of protein-surface interaction. This protein-

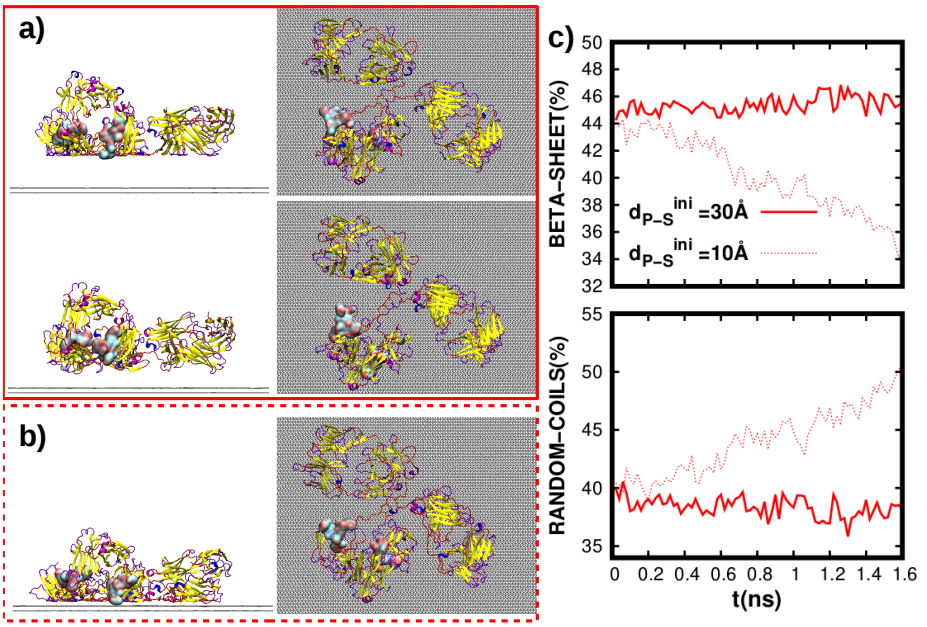



**Figure 2.8:** ***Analysis of the structural changes that can be induced during the SMD simulation for restraining the $C_\alpha$ belonging to 16 cysteine residues. a)*** *We here show both side (left) and top (right) views of the initial (top) and final (bottom) configurations obtained after a SMD simulation of the IgG adsorption process with an initial protein-surface distance larger than the one used in our adsorption protocol ($d_{P-S}^{ini} = 30$ Å $> 10$Å, see sec. 2.2.1). The solvent method used for this simulation is the OBC implicit solvent method [55]. The color representation used for the protein-substrate system is the same as in Fig. 2.2. For comparison, we also include in* ***b)*** *the final configuration obtained after the SMD process used in the adsorption protocol, i.e. starting at $d_{P-S}^{ini} = 10$ Å. In* ***c)****, the time evolution of the $\beta$-sheet (top) and random-coils (bottom) content of the IgG during both SMD processes, i.e. starting at $d_{P-S}^{ini} = 30$ Å (continuous line) and at $d_{P-S}^{ini} = 10$ Å (dashed line) is also shown.*

structural-instability promoted by the contact is in agreement with data shown in Fig. 2.4, and shall be subject of a thorough discussion in following sections.

In summary, upon analysis of the IgG tertiary and secondary structure shown in Figs. 2.4, 2.6 and 2.7, we may conclude that OBC and HCT methods predict a major structural rearrangement of the IgG upon adsorption to graphene. Altogether this is incompatible [127] with the preservation of both structure and

bio-functionality shown by the experimental observations [56]. Therefore, our results firmly confirm that improvements on OBC and HCT implicit methods are required in order to correctly describe the adsorption process of the IgG.

## 2.4 Effect of $E_{SOL(nonel)}$ in the adsorption process

In ref. 97, similar results were found for a completely different protein, the BSA. Antibodies such as IgG and serum albumin are different in size, secondary structure composition and net charge. Despite these differences, implicit solvent simulations predicted also a major BSA unfolding upon its adsorption [97], contrasting with both explicit solvent simulations [97] and recent experimental reports [128]. Therefore, the observed IgG unfolding upon adsorption in implicit solvent simulations seems not to be a specific feature of this protein, but a general characteristic of implicit solvent methods. The authors in ref. 97 speculated that such degree of unfolding could be attributed to a strong energy gradient that arises from an abrupt change of the non-electrostatic solvation contribution ($E_{SOL(nonel)}$) when the protein contacts the surface. In this section we test the validity of this assumption through a detailed energy decomposition analysis.

In Fig. 2.9a, we plot the energetic change of the protein-substrate system ($\Delta E^T$) as a function of the distance between them ($d_{P-S}$) using the HCT method. To compute it, we initially placed the IgG at $d_{P-S} \sim$ 12.5 Å and brought it closer to the surface in steps of 0.1 Å , see Fig. 2.9b. The total energetic change from its initial value, i.e $E^T$ at $d_{P-S} \sim$ 12.5 Å, is evaluated in all intermediate steps of $d_{P-S}$. We also compute the different contributions to the total energy change:

$$\Delta E^T = \Delta E^{P-S}_{vdW} + \Delta E_{SOL(el)} + \Delta E_{SOL(nonel)} \ , \tag{2.8}$$

where $E^{P-S}_{vdW}$ is the graphene-protein vdW interaction energy, $E_{SOL(el)}$ is the electrostatic component of the solvation energy and $E_{SOL(nonel)}$ the non-electrostatic contribution to the solvation energy (see eqs. 2.1 and 2.2). In Fig. 2.9a, we observe that the change of the total energy of the system ($\Delta\ E^T$) strongly depends on the protein-surface separation. From here on, we focus

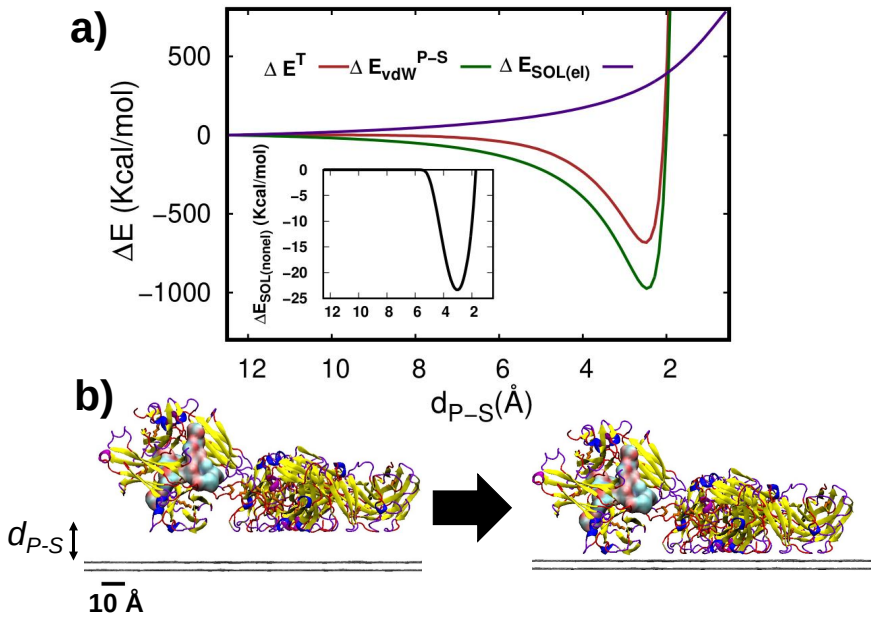



**Figure 2.9:** ***Description of the*** $\Delta E(d_{P-S})$ ***with implicit solvent methods.*** ***a)****Change of the total interaction energy (brown) with the protein-surface distance* ($d_{P-S}$) *using the HCT implicit [54] solvent method. The contributions to this change are the protein-substrate interaction (green), the electrostatic solvation contribution (purple) and the non-electrostatic solvation contribution (inset, in black color). The distance* $d_{P-S}$ *is defined as the z coordinate of the lowest hydrogen atom of the proline 889, considering that the upper graphene layer is at z=0.* ***b)****Representation of the system configuration used to compute this energy change. The color representation is the same as in Fig. 2.2.*

only its evolution at $d_{P-S} \geq$2.5 Å where the adsorption process is energetically favorable ($\Delta E^T$ decreases). It is important to remark that in these calculations, the coordinates of the protein were kept fixed in all steps, i.e. not allowing for the structural reorganization of the protein. The protein configuration used is the final adsorption configuration obtained in explicit solvent simulations, as it is the folded configuration that leads to the largest SASA change when the protein contacts the surface.

First we consider the change of the non-electrostatic solvation contribution ($\Delta E_{SOL(nonel)}$) as a function of $d_{P-S}$, see inset of Fig. 2.9a. From here we

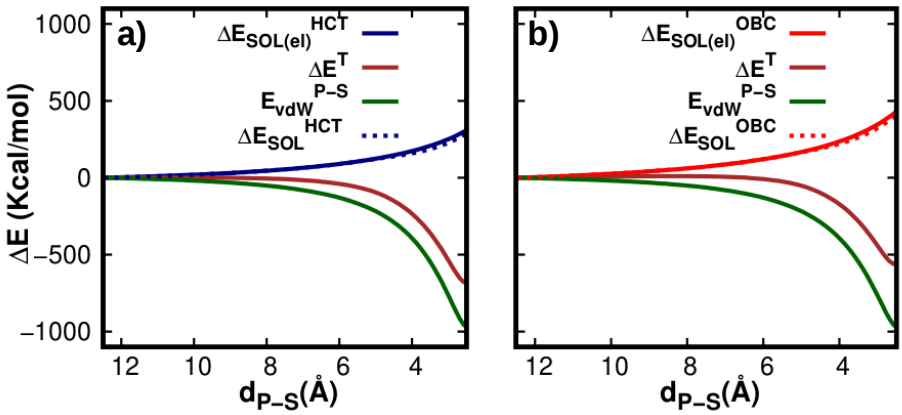



**Figure 2.10:** ***Change of the electrostatic component of the solvation energy, $E_{SOL(el)}$, of the protein+substrate system with the distance between them ($d_{P-S}$).*** *Here the results obtained using the* ***a)****HCT [54] and* ***b)****OBC [55] implicit solvent methods are showed. The change of the total energy of the system ($\Delta E^T$, in brown),the change of the protein-substrate interaction ($\Delta E^{P-S}_{vdW}$, green) and the total solvation contribution ($E_{SOL} = \Delta E_{SOL(el)} + \Delta E_{SOL(nonel)}$, dashed line) are also shown for comparison.*

observe not only that $E_{SOL(nonel)}$ changes smoothly with $d_{P-S}$ but also that its contribution to the total energy change is negligible, i.e. $\Delta E_{SOL(nonel)} = 3.5\%\Delta E^T$. These two aspects support that $E_{SOL(nonel)}$ plays a minor role, if at all, in the adsorption process. Therefore, $\Delta E_{SOL(nonel)}$ alone cannot account for the spurious protein unfolding observed in implicit solvent simulations [92, 97], showing that the previous conjecture [97] was wrong. Note that this negligible contribution of $\Delta E_{SOL(nonel)}$ to the total energy change happens in a simulation designed to maximize it, i.e. we picked the IgG folded structure with the largest CSA which should result in the largest $\Delta E_{SOL(nonel)} = \gamma^{\star} CSA$.

We then focus on the change of the other energy components: $E^{P-S}_{vdW}$ and $E_{SOL(el)}$. Firstly, we consider the vdW protein-surface interaction ($E^{P-S}_{vdW}$) evolution. From Fig. 2.9a, we observe that this energy component accounts for most of the variation in the total energetic change as it decreases by $E^{P-S}_{vdW} \sim 1000$ kcal/mol. Thus, this attractive interaction is the main driving force of this process. Secondly, we consider the variation of the $E_{SOL(el)}$, which, for protein adsorption on

neutral surfaces, represents the solvent electrostatic screening within the protein. This component is the only difference between OBC and HCT implicit solvent descriptions, see sec. 2.2.2, and its change when the protein approaches to the surface is represented in Fig. 2.10. From this figure, we extract two important features about this energetic component. First, the OBC implicit solvent method predicts a larger $E_{SOL(el)}$ variation than the HCT implicit solvent method, see Fig. 2.10a-b. This result may explain the slight differences observed in the adsorption configurations obtained with these two implicit solvent methods. Second, we see that $\Delta E_{SOL(el)}$ is contributing significantly to the total energy change, namely by 30%. In particular, for HCT(OBC) we observe that $E_{SOL(el)}$ increases by 300(400) $\mathrm{kcal\ mol^{-1}}$ during the adsorption, which highlights the highly repulsive nature of this interaction, see Fig. 2.10. Additionally, this variation shows that, when approaching the surface, the protein is losing a significant amount of its electrostatic screening, what will certainly affect its internal stability. In the following sections, we are going to use this information to propose a new hypothesis for explaining the observed unfolding of the IgG upon adsorption.

## 2.5 vdW interaction: implicit vs explicit solvation description

### 2.5.1 vdW protein-substrate interaction: comparison between explicit and implicit solvent results

In this section, we focus on the evolution of all the vdW energetic components of the system during the adsorption process to understand how they change with the solvation model and their role in the IgG adsorption. We first analyze the time evolution of the vdW protein-substrate interaction, i.e. the variation of that magnitude during the adsorption time ($\Delta_t E_{vdW}^{P-S}$), for all the solvents here considered, see Fig. 2.11a. Note that henceforth the results presented always refer to the MD simulations characterized in sec. 2.3. From Fig. 2.11a, we observe that in all cases $\Delta_t E_{vdW}^{P-S} < 0$, i.e. the vdW protein-substrate interaction favors the adsorption process regardless of the solvent method used. In contrast with this similarity, two major differences may also be noticed when comparing

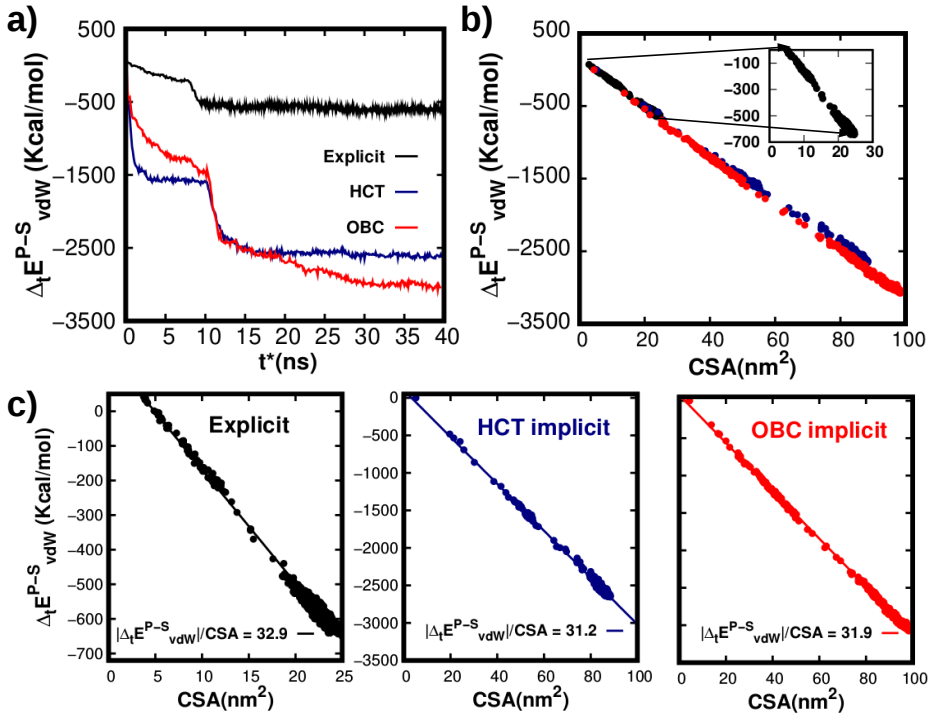



**Figure 2.11:** ***Evolution of the interaction energy between the protein and the substrate,*** $E_{vdW}^{P-S}$***. a)****Variation of the interaction energy between the protein and the substrate during the simulation time for the whole adsorption process (*$\Delta_t E_{vdW}^{P-S}$*). The obtained results using the TIP3P explicit solvent [114] (black), the HCT implicit [54] solvent (blue) and the OBC implicit [55] solvent (red) are all showed. The time of simulation for the explicit solvent case has been rescaled in all the figures using the relation* $t* = t_{sim}/4.6$*. **b)****Evolution of the interaction energy between the protein and the substrate with the contact surface area (CSA) for the three solvent method used: Explicit [114] (black), HCT implicit [54] (blue) and OBC implicit [55] (red). The evolution of* $\Delta_t E_{vdW}^{P-S}$ *with CSA using explicit solvent has been also represented in the inset. **c)****Rate of change of* $E_{vdW}^{P-S}$ *with the CSA for the three solvent methods here used. This rate of change has been computed via a linear regression of the MD data.*

explicit and implicit solvent results. First, at the end of the adsorption process, we observe that $|\Delta_t E_{vdW}^{P-S}|$ is $\sim 5$ times larger in implicit solvent simulations as compared to the explicit solvent ones. It is worth to mention that, among implicit solvent methods, the largest energy is obtained with the OBC method,

which is also the one where a larger protein unfolding is observed (see Figs. 2.6 and 2.7). The second major difference between implicit and explicit solvent results concerns the variation rate of $E_{vdW}^{P-S}$ in the initial 10 ns stage prior to the enhanced adsorption protocol. In explicit solvent simulations we observe a smooth decrease of $E_{vdW}^{P-S}$ as the protein approaches the surface. This is at odds with OBC/HCT results, where we observe an initial abrupt change of this interaction. In fact, in explicit solvent simulations at the end of the whole adsorption process ($\Delta t = 150$ ns) the total $\Delta_t E_{vdW}^{P-S}$ is much smaller than the corresponding change during the initial 10 ns of implicit solvent simulations.

In order to understand the origin of the different $\Delta_t E_{vdW}^{P-S}$ obtained at the end of each simulation we now analyze how this interaction correlates with the CSA. From Fig. 2.11b, we observe that $\Delta_t E_{vdW}^{P-S}$ depends linearly on CSA, with a similar slope for all the solvent methods here considered. Moreover, when computing the slope we obtain $\Delta_t E_{vdW}^{P-S}/\Delta CSA \sim 32\ kcal\ mol^{-1}\ nm^{-2}$, see Fig. 2.11c, which is in agreement with vdW interactions energies per unit area obtained in gold-standard quantum chemistry simulations (10-100 $kcal\ mol^{-1}\ nm^{-2}$) [129]. Besides the obvious conclusion that regardless the solvent method used the interaction $E_{vdW}^{P-S}$ is computed similarly, this result highlights that implicit solvent methods predict a larger $\Delta_t E_{vdW}^{P-S}$ only because the CSA is also larger at the end of the adsorption process. Thus, larger $\Delta_t E_{vdW}^{P-S}$ is solely a consequence, and not the origin, of why different solvation methods result in very different final adsorption configurations. Additionally, this indicates that, to trace the origin of these differences, we must understand why CSA increases more rapidly in implicit solvent methods as compared to explicit ones.

### 2.5.2 CSA evolution: origin of the differences observed in explicit and implicit solvent simulations

The slower increase rate of CSA obtained in explicit solvent simulations can be understood in light of the time evolution of all intermolecular vdW interactions ($\Delta_t E_{vdW}$) shown in Fig. 2.12a. There, we observe that, in the initial free adsorption regime (t$<10$ ns), the attractive protein-substrate interaction energy ($E_{vdW}^{P-S}$) is compensated by the sum of two repulsive solvent

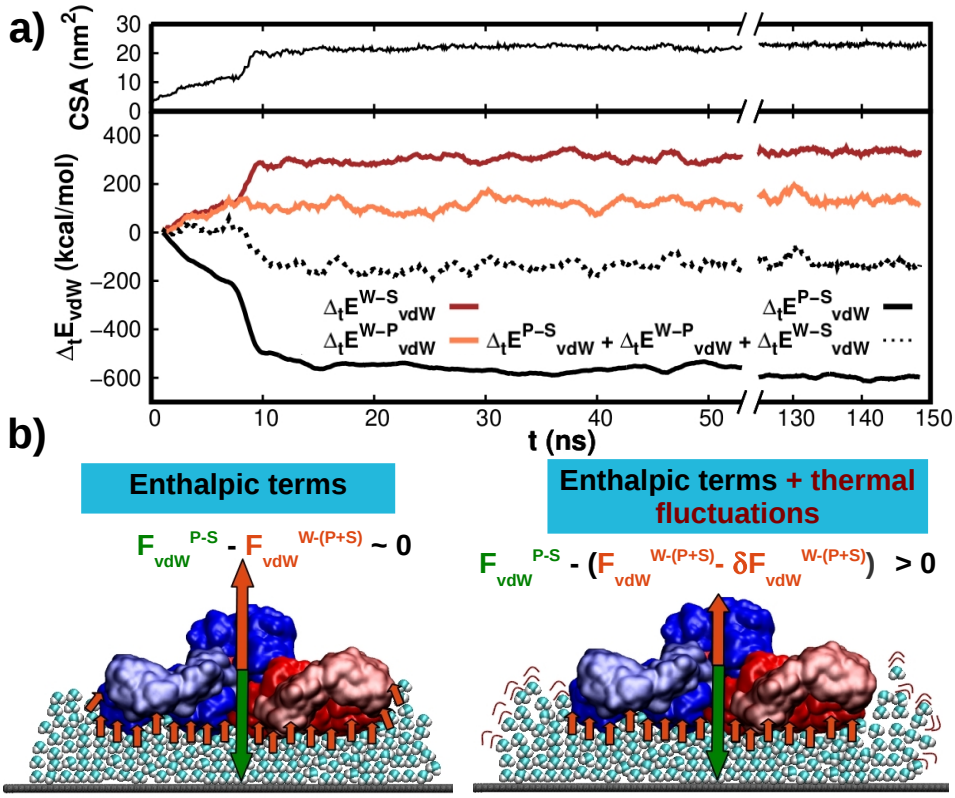



**Figure 2.12:** ***Role of the vdW interaction in explicit solvent simulations. a)*** *Time evolutions of the CSA (top panel) and of the sum of all intermolecular vdW interaction energies (bottom panel, black dashed line) using explicit solvent. These interaction energies are the vdW water-protein interaction,* $E_{vdW}^{W-P}$ *(orange), the vdW water-substrate interaction,* $E_{vdW}^{W-S}$ *(brown), and the vdW protein-substrate interaction,* $E_{vdW}^{P-S}$ *(black).* ***b)*** *The enthalpic vdW energy balance between the attractive protein-surface interaction and the force opposing solvent squeezing-out is broken by the water thermal fluctuations, that lead to transient force unbalances and, ultimately, to protein adsorption.*

interactions energies ($E_{vdW}^{W-(P+S)} = E_{vdW}^{W-S} + E_{vdW}^{W-P}$). This can be thought as an energy balance where the energy gained in adsorbing the protein is expended in breaking the hydration layers separating it from the surface and thus $\Delta_t E_{vdW}^{P-S} + \Delta_t E_{vdW}^{W-(P+S)} \sim 0$. As a result, it follows that the the vdW component of the effective adsorption force is zero, i.e. $F_{vdW}^{eff} = 0$. This result leads to an apparent contradiction as, on one hand, the adsorption is spontaneous (as

shown by the CSA in Fig. 2.12a) and, on the other hand, the enthalpic vdW energy balance is null. To solve this paradox, we must now consider the entropic contribution to the adsorption energy. Although we do not explicitly quantify this contribution, it is considered in our MD simulations, as these simulations sample the whole phase space of the system. The thermal vibrations of the water molecules allow for a protein to diffuse while in solution. Additionally, given the isotropic nature of the medium, statistically no direction is privileged with respect to other in this diffusion. If we now consider a case where an adsorbing substrate breaks this isotropy, we may argue that this diffusion will privilege displacements towards the surface even if the enthalpic energy gain is null. The mechanism by which this occurs can be thought as follows. Let us consider a protein surrounded by water molecules in the vicinity of a substrate as shown in Fig. 2.12b. The force felt by the protein is a sum of an attractive interaction with the surface, i.e. $F_{vdW}^{P-S}$, with a force opposing solvent squeezing-out , i.e. $F_{vdW}^{W-(P+S)}$, necessary for the adsorption to occur. Although our MD simulations indicate that, on average, these two compensate each other, the solvent thermal fluctuations lead to temporary force unbalances. As schematically shown in Fig. 2.12b, these unbalances can lead to a squeeze-out of the solvent and subsequent approximation of the protein towards the substrate. This analysis implies that the adsorption process using explicit solvent happens as a consequence of a subtle effect, i.e. the slight separation of water molecules from the protein surface due to thermal oscillations. Therefore, a slow approach of the protein to the surface is expected when using explicit solvent from which it follows a slow evolution of the CSA.

In order to understand why in implicit solvent MD simulations the CSA evolution is much faster than in explicit solvent MDs, we focus in the vdW solute-solvent interactions of the former. As mentioned in sec. 2.4, these interactions are included in implicit solvent simulations via the non-electrostatic solvation component, i.e $E_{SOL(nonel)}$, whose contribution to the total energy is negligible as compared to the $E_{vdW}^{P-S}$. This is at odds with the fact that vdW interactions are non-specific. From this property, it follows that the vdW interaction force with its surrounding should be about the same either if it is in contact with the surface or with the solvent, assuming that the total area enclosing the protein is constant during the process. Although this does seem to be the case in explicit solvent simulations, this is not so for implicit solvent since $\Delta E_{SOL(nonel)} + \Delta E_{vdW}^{P-S} \sim \Delta E_{vdW}^{P-S} \neq 0$ as shown in sec. 2.4. Therefore, $E_{SOL(nonel)}$ is unable to describe the important vdW solvent contribution to the total energy during the adsorption process, which could be associated to processes such as the cost of breaking the solvation shells surrounding the protein and the substrate. Moreover, this also results in an overestimation of the effective adsorption force in implicit solvent simulations as $F_{vdW}^{W-(P+S)} \sim 0$ and thus $F_{vdW}^{eff} \sim F_{vdW}^{P-S}$. This overestimation has two direct consequences: a quasi-instantaneous adsorption process, and a much higher adsorption force as compared with explicit solvent simulations. This two consequences allow us to explain the two differences observed between implicit and explicit solvent results in the CSA evolution, i.e. faster evolution and the larger final absolute values.

## 2.6 Electrostatic interaction: implicit vs explicit solvation description

### 2.6.1 Evolution of the electrostatic component of the solvation energy in explicit and implicit solvent simulations

The previous section highlights the importance of the vdW energy component of the solvation energy in compensating the large protein-surface interaction. When it is underestimated (as in the implicit solvent simulations), the CSA between the protein and the surface dramatically increases, see Fig. 2.11a. Although

this rapid change of the CSA certainly affects the protein stability, it does not suffice to understand the unfolding observed in implicit solvent simulations. In this section, we focus on the other component of the system solvation energy, i.e. the electrostatic component ($\Delta_t E_{el}^{SOL}$), to completely understand that IgG unfolding. This energy contribution is defined as the difference between the electrostatic energy of the protein-substrate system immersed in water with respect to its electrostatic energy in vacuum [115]. In implicit solvent simulations this energy component is defined as the GB solvation (see eq. 2.2), i.e.

$$E_{el}^{SOL,implicit} = E_{SOL(el)} = E_{GB} \tag{2.9}$$

For our system, as the graphene atoms are not charged, this term accounts for the solvent screening of the protein intra-molecular electrostatic interactions. In explicit solvent simulations, the definition of $E_{el}^{SOL}$ [115] results in the following expression:

$$E_{el}^{SOL,explicit} = E_{el}^{W} + E_{el}^{W-P} + E_{el}^{W-S} \tag{2.10}$$

where $E_{el}^{W}$ accounts for the electrostatic internal energy of the water, $E_{el}^{W-P}$ accounts for the electrostatic water-protein interaction and $E_{el}^{W-S}$ accounts for the electrostatic water-surface interaction. Note that as the graphene atoms are not charged, $E_{el}^{W-S} = 0$.

In Fig. 2.13a, we show the time evolution of $\Delta_t E_{el}^{SOL}$ for both explicit and implicit solvent methods. From that figure, we observe that, in explicit solvent simulations, the electrostatic contribution of the solvent energy hardly changes during the adsorption process as $\Delta_t E_{el}^{SOL} \sim 0$. To analyze in detail this result, we have decomposed in Fig. 2.13b the evolution of $E_{el}^{SOL}$ into its different non-null components, i.e $E_{el}^{W}$ and $E_{el}^{W-P}$. From this figure, we may derive two different results. Firstly, the electrostatic protein-water interaction does not favor the adsorption as $\Delta_t E_{el}^{W-P} > 0$. Secondly, $\Delta_t E_{el}^{W-P}$ is strongly correlated with $\Delta_t E_{el}^{W}$, as during the whole adsorption process $\Delta E_{el}^{W} \sim -\Delta E_{el}^{W-P}$. These two results indicates that while the protein adsorbs, the loss of water-protein electrostatic energy arising from protein's dehydration is fully compensated by the gain of water's internal electrostatic energy. This internal energy gain is in turn a result of new water-water h-bonds being formed during protein's dehy-

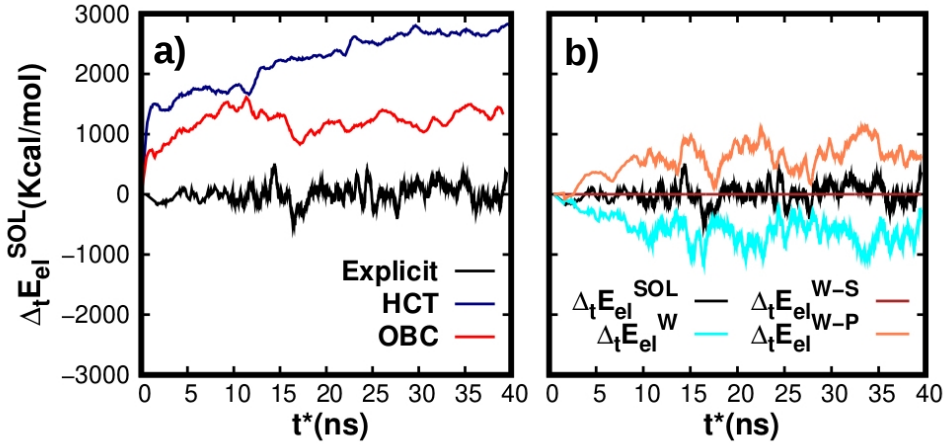



**Figure 2.13:** ***Variation of the electrostatic contribution to the solvation energy during our MD simulations ($\Delta_t E_{el}^{SOL}$).*** ***a)****Using the three solvent methods here considered: HCT implicit [54] solvent (blue), OBC implicit [55] solvent (red) and explicit solvent [114] (black)* ***b)****Using explicit solvent (black line), with the contribution of each of the energy components of $E_{el}^{SOL}$: $E_{el}^{W}$ (cyan), $E_{el}^{W-P}$ (orange) and $E_{el}^{W-S}$ (brown)*

dration. The balance between these two mechanisms explains why $E_{el}^{SOL}$ does not seem to contribute enthalpically to the adsorption process.

In sec. 2.5, we showed that, in implicit solvent simulations, the protein adsorbs instantly due to a high protein-surface vdW interaction. This fast adsorption results in a rapid increase of the CSA during the first 2.5 ns for both implicit solvent methods, see Fig. 2.14a. Additionally, in Fig. 2.14b we also observe that, during this time range, $\Delta_t E_{el}^{SOL} > 700$ kcal/mol, which shows that the electrostatic contribution of the solvent energy is working against the protein adsorption thus acting as an effective repulsive force. However, this *repulsive* force is unable to hinder the protein adsorption as the attractive vdW protein-surface interaction compensates this energy loss, see Fig. 2.9a. Still, as a result of the adsorption the system is left with a large electrostatic energy unbalance which may only be alleviated by a major structural rearrangement of the charges present on the system, which in implicit solvent are only the protein atoms. Remarkably, in Fig. 2.14c we observe that the protein's internal electrostatic energy,

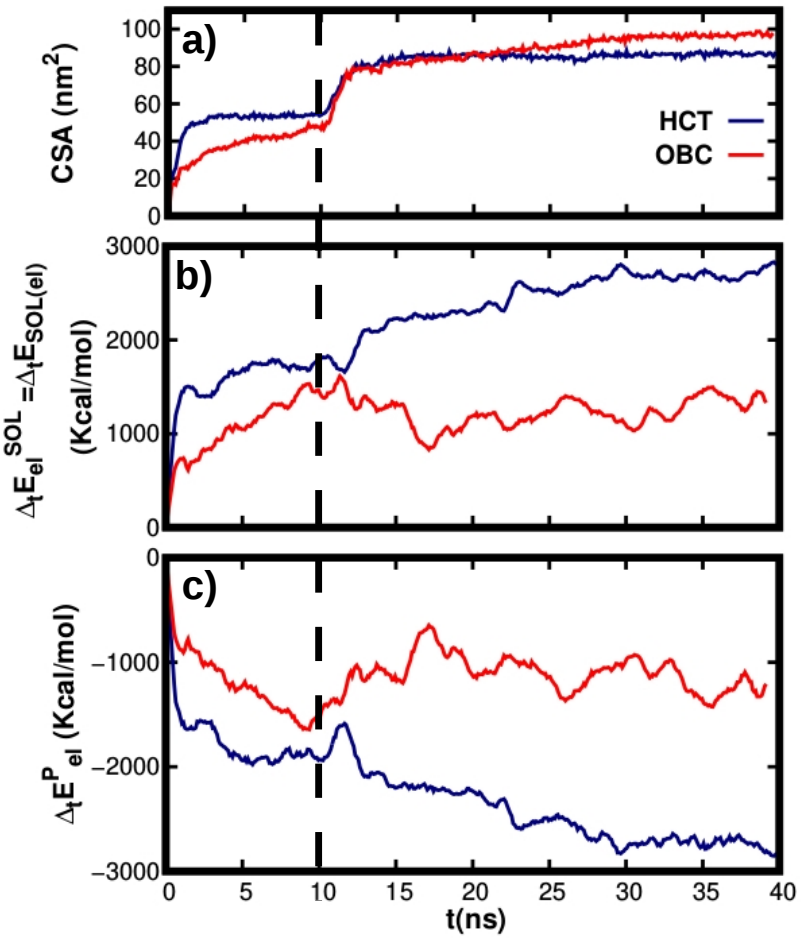



**Figure 2.14:** ***Correlation between the solvent electrostatic screening and the internal electrostatical energy of the protein.*** *Time evolution during the whole adsorption process of* ***a)****the CSA,* ***b)****the electrostatic component of the solvation energy,* $E_{SOL(el)}$*, and* ***c)****the internal electrostatic energy of the protein,* $E_{el}^{P}$ *using the HCT [54] (blue) and the OBC [55] (red) implicit solvent methods.*

$E_{el}^{P}$, not only decreases significantly as a result of the adsorption process, but also the variation of this energy component is strongly correlated with $\Delta_t E_{el}^{SOL}$. This correlation demonstrates that the aforementioned electrostatic unbalance leads to a major protein structural rearrangement. Moreover, it is important to note that all these energy changes (namely $\Delta_t E_{el}^{P}$, $\Delta_t E_{el}^{SOL}$ and $\Delta E_{vdW}^{P-S}$)

occur almost instantaneously, i.e. within the first two nanoseconds. The large energy unbalances, coupled with their fast time rates, result in a swift protein unfolding upon adsorption when using implicit solvent methods, as clearly evidenced by its secondary structure evolution shown in Fig. 2.7a. Therefore, on the overall, the protein adsorption using implicit solvation methods can be seen as a three step process. Firstly the protein quickly approaches to the surface as a consequence of the vdW interaction. Secondly, as a consequence of the protein-surface approach, a large electrostatic repulsive force emerges from the solvent contribution. Lastly, to reduce this energy unbalance, the protein is *instantaneously* forced to rearrange its atoms, which in turn leads to the protein's unfolding. As a result, one may argue that this spurious unfolding observed in many different implicit solvent MD works concerning the protein adsorption to uncharged surfaces may be understood in terms of the ill-characterization of these three elements.

### 2.6.2 Gain of water-water electrostatic interaction: only considered in explicit solvent simulations

One of the two major differences observed between implicit and explicit solvent MD simulations concerns the evolution of the electrostatic energy component of the solvation energy. In particular, Fig. 2.13a shows that $\Delta_t E_{el}^{SOL,implicit} > 0$ whilst $\Delta_t E_{el}^{SOL,explicit} \sim 0$. This major qualitative and quantitative difference may be traced back to the definition of this term in the two solvent methods, see eqs. 2.9 and 2.10. In the explicit solvent simulation $\Delta_t E_{el}^{SOL,explicit}$ considers two different solvent effects: the changes in protein's intra-molecular solvent screening via the $\Delta_t E_{el}^{W-P}$ term, and the changes in the solvent-solvent electrostatic interaction via the $\Delta_t E_{el}^{W}$ term. Notably, in Fig. 2.13b, we observed that these two terms are highly correlated. This correlation highlights the feedback mechanism in which electrostatic energy losses arising from protein's dehydration are compensated by similar energy gains in the solvent (through e.g. additional solvent-solvent h-bonds). This loss(gain) of electrostatic interactions of the water molecules with protein(water) is what ultimately regulates the electrostatic contribution of the solvation energy, leading to $\Delta_t E_{el}^{SOL,explicit} \sim 0$ during ad-

sorption. However, in implicit solvent simulations $\Delta_t E_{el}^{SOL,implicit}$ only accounts for the first of these two effects, i.e. the loss of the solvent screening of the intramolecular protein electrostatic interactions. For this reason, using this solvent method, the electrostatic contribution to the solvent energy suppose a positive energy gradient ($\Delta_t E_{el}^{SOL,implicit} > 0$), which only can be balanced with the reorganization of the protein's atoms. This difference remarks the importance of considering the gain of the water-water electrostatic interaction during adsorption, as only when it is omitted (implicit solvent), the evolution of $E_{el}^{SOL}$ affects the structural stability of the protein.

### 2.6.3 Loss of water-protein electrostatic interaction: different description in implicit and explicit solvation methods

The other difference between implicit and explicit solvent simulations concerns the loss of water-protein electrostatic interaction while the protein adsorbs/dehydrates. To understand this difference, we first need to comprehend how $E_{el}^{SOL,implicit}$ describe this process. To this aim, we have analyzed the variation of $E_{el}^{SOL,implicit}$ when a charged model molecule approaches to a neutral model surface. Both systems, i.e the molecule and the surface, are conformed by spheres/atoms of radius 1 $R_{vdW}$, as shown in the inset of Fig. 2.15. The spheres of the molecule are equally positively charged, while the ones of the surface are neutral. We have used the same definition of $E_{el}^{SOL,implicit} = E_{SOL(el)} = E_{GB}$ as in the HCT method [54], see sec. 2.2.2. According to that definition, each charged sphere/atom $i$ of the solute contributes to the total electrostatic component of the solvent energy via the expression:

$$\begin{aligned} E_{SOL(el)} &= \sum_i E_{SOL(el)}(i) \\ &= -(\frac{\epsilon_\omega - 1}{2\epsilon_\omega}) \sum_i \sum_j \frac{q_i q_j}{r_{ij} \cdot \sqrt{1 + \frac{R_i^{born} R_j^{born}}{r_{ij}^2} e^{-r_{ij}^2/4R_i^{born}R_j^{born}}}}, \end{aligned} \tag{2.11}$$

where $R_i^{born}$, i.e. the effective Born radii, reflects the degree of burial of the sphere/atom $i$ inside the molecule. We have computed numerically that expression to analyze how changes $E_{SOL(el)}$ with the molecule-surface distance,$d_{M-S}$.

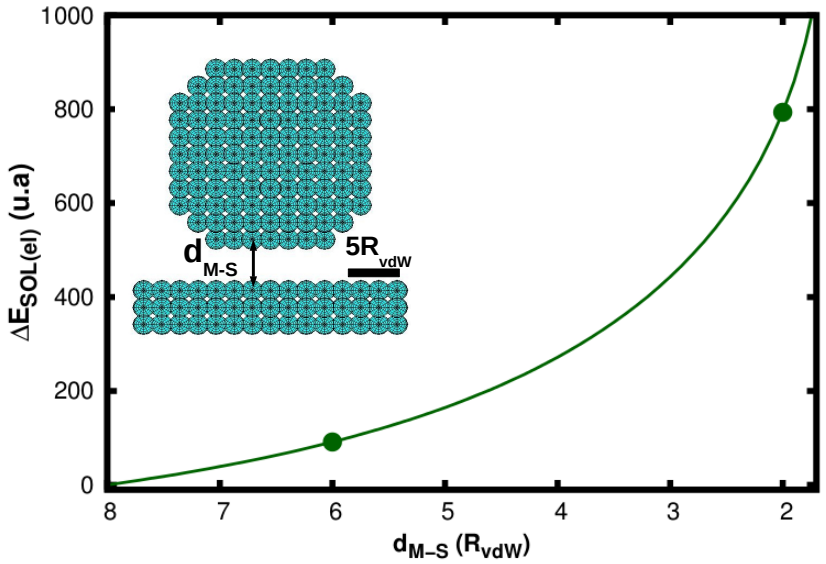



**Figure 2.15:** ***Modeling the loss of the solvent electrostatic screening predicted by implicit solvent methods.*** *Change of the electrostatic component of the solvent energy, $E_{SOL(el)}$, when the distance between the model molecule and surface ($d_{M-S}$) decreases. The model molecule is conformed by 852 positively charged spheres and the surface is formed by 667 neutral spheres distributed in three layers. The radius of all the spheres is 1 $R_{vdW}$. A representation of the whole system is showed in the inset.*

In Fig. 2.15 we observe that, as the molecule approaches to the surface, $E_{SOL(el)}$ largely increases. This result is very similar to the one showed in Fig. 2.9a, i.e. when the molecule and the surface are the IgG protein and the graphene double layer, respectively. This similarity highlights that the observed increase of $E_{SOL(el)}$ during adsorption is not a specific consequence of the adsorption process of the IgG, but a general feature of the implicit solvent methods. To understand this feature, we analyze the contribution of each molecule's sphere to $\Delta E_{SOL(el)}$ for different $d_{M-S}$ values. In Fig. 2.16a, we observe an enhancement of the contribution of the spheres closer to the surface when $d_{M-S}$ decreases. This enhancement, in accordance to eq. 2.11, arises from an increase of the Born radii of these "close to surface" spheres as shown in Fig. 2.16b Therefore, the increase of $E_{SOL(el)}$ when a charged molecule approaches to a

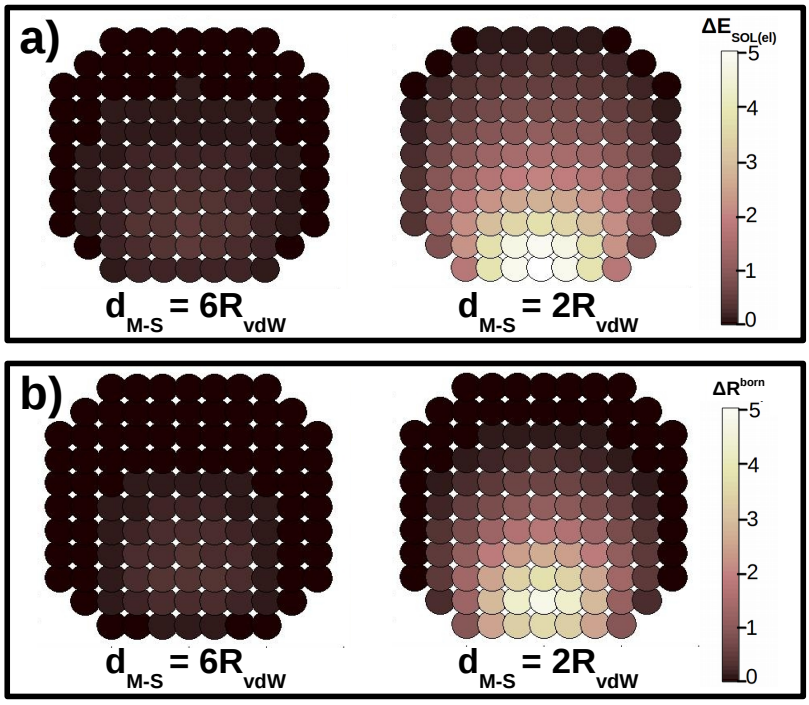



**Figure 2.16:** ***Change of the*** $\Delta E_{SOL(el)}$ ***per sphere. a)***$\Delta E_{SOL(el)}$ *per sphere when* $d_{M-S} = 6\ R_{vdW}$ *(left) and* $d_{M-S} = 2\ R_{vdW}$ *(right).* ***b)****Change of the Born radii* ($R_{born}$) *[54] per sphere while decreasing the distance between the model molecule and the model surface* ($d_{M-S}$)*. The results here shown are also for* $d_{M-S} = 6\ R_{vdW}$ *(left) and* $d_{M-S} = 2\ R_{vdW}$ *(right).*

neutral surface accounts for the following process. When the molecule is far from the surface, the degree of burial of one molecule sphere, only depends on its position inside the molecule. For this reason, these "close to surface" spheres/atoms are outer spheres in that situation (small Born radii). However, when the molecule approaches to the surface, the position of the surface spheres starts to affect the degree of burial of the molecule spheres. As a consequence, these "close to surface" spheres/atoms start to act as inner spheres, i.e their Born radii enlarges. This ultimately leads to the increase of $E_{SOL(el)}$, explaining how the loss of water-protein electrostatic interactions is described in implicit solvent simulations.

In explicit solvent simulation, the loss of water-protein electrostatic interactions is described via the removal of the water molecules that are in the region between the molecule and the surface. This removal process is more gradual than the process of change of the Born radii used in implicit solvent simulations. Additionally, as implicit solvent methods overestimate the effective vdW protein-surface attractive force, the adsorption process is faster in these simulations than using explicit solvent, see sec. 2.5. These two factors cause, as shown in Fig. 2.13, a larger and faster loss of water-protein electrostatic interaction in implicit solvent simulation ($\Delta_t E_{el}^{SOL,implicit}$) than in explicit solvent simulations ($\Delta_t E_{el}^{W-P}$). Considering the direct correlation between this loss and the reorganization of the protein's atoms, this difference is translated into instantaneous and great changes in the structure of the IgG protein. Therefore, the abrupt loss of water-protein electrostatic interactions contributes also to the unfolding of the protein in implicit solvent simulations.

## 2.7 Conclusions

In this chapter, the IgG adsorption process on a graphene surface was studied using both explicit and implicit solvent MD simulations. The results of the explicit solvent simulation show that water molecules play an active role in the adsorption process. When the protein approaches the graphene surface, the loss of water-protein electrostatic interaction becomes balanced by the gain of the water-water electrostatic interaction. This regulates the electrostatic contribution of the solvation energy. Additionally, the cost of breaking protein and surface solvation shells, compensates the vdW attraction that exists between them. Despite of this energy compensation, the protein adsorbs on the surface. This indicates that the driving force of the adsorption process does not only present an enthalpic origin. The thermal vibrations of the water molecules allows the protein to diffuse, and the presence of an attractive protein-surface force guides this protein diffusion towards the surface. Therefore, entropic effects seems to be important in the adsorption process of a protein on a hydrophobic surface.

In contrast, implicit solvent methods underestimate the cost of breaking the

solvation shells of the protein and the substrate. This leads to an overestimation of the effective vdW protein-surface attractive force, causing an instantaneous adsorption process. In addition to the fast adsorption process, a large energy gradient arises from the loss of the water-protein electrostatic interaction. Contrary to explicit solvent simulation, that energy gradient is not balanced by the gain of water-water electrostatic interaction, which is not included in implicit solvent simulations. Therefore, the only way of alleviating that energy unbalance is by reorganizing the position of the only charges of the system, the protein atoms. This reorganization happens abruptly, which leads to the unfolding of the protein. Thus, with these MD simulations, we have been able to understand the factors that lead to unfolded protein configuration upon adsorption when implicit solvent methods are used. To prevent the unfolding of the protein, we propose to improve both non-electrostatic and electrostatic components of the solvation energy. Concerning the first, we think that, in this current formulation, it underestimates the vdW repulsive solute-solvent interaction, which according to explicit solvent simulations has an important role in the adsorption process. Regarding the electrostatic component, we propose that a correction in the Born radii definition is needed in order to obtain a more gradual loss of the solvent screening of the intra-protein electrostatic interactions.

# 3 Tuning the Mechanical Response of ssDNA films via its hydration level

The influence of water on a vast range of processes occurring in biological systems, such as protein folding or DNA hybridization, could be use to tune in a controlled manner the properties of such systems. Experiments [130] have recently shown that the mechanical response of a biosensing device conformed by a SAM of ssDNA strands can be directly controlled with its hydration. In fact, they are able to tune this mechanical response from an upward to a downward deflection only by increasing the relative humidity of the system. In order to understand the mechanisms behind this water-induced change, in this chapter we have studied with atomistic MD simulations the variation of the mechanical response of a high-densely packed SAM with the hydration level. Our MD simulation results reproduce the hydration-driven variation of the SAM mechanical response observed experimentally for two different values of the SAM grafting density. What's more, we have been able to describe in detail the physical origin of the two main trends of the SAM mechanical response: upward, and downward deflection. First, at low hydration levels, the main driving force of the SAM mechanical behaviour is the water-ssDNA interaction. Consequently, the number of water-ssDNA h-bond bridges, i.e. formed between a single water molecule and several ssDNA atoms, increases with the hydration. This stiffens the ssDNA structure as the hydration increases, causing a more energetic (larger) deflection of the SAM with the hydration (upward deflection). However, when the water-water interaction becomes relevant to the SAM behaviour, the ssDNAs tend to enlarge the volume of the nanopores by increasing their height. That reduces

the maximum curvature angle between two consecutive ssDNAs, hindering the increase of the SAM curvature (downward deflection).

## 3.1 Introduction

### 3.1.1 Water interaction with polymer monolayers

The interaction of water with different interfaces at the nanoscale has attracted much attention in multiple areas of science and technology, such as sensing, catalysis, tribology and protein adhesion. In the last decade, a great deal of effort has been devoted to understand the properties of this interaction both when the water molecules are static [131–134] and when they are in motion [135–138] . In these two contexts, atomistic MD simulations have been used as an essential tool to shed light on the water-interface interaction properties [133, 136, 138, 139] , as they accurately reproduce the experimental behavior of water molecules in contact with interfaces [113, 133] .

Recently, a new class of biointerfaces has started to be a subject of growing interest [140–149] . These biointerfaces consist on densely end-grafted polymer assemblies and are commonly known as *polymer brushes*. The possible applications for these polymer assemblages range from lubrication [143, 150, 151] to the reduction of protein adsorption [146, 152] . Consequently, these systems has been widely studied during the past years, with the aim of controlling their tribological, mechanical and structural properties. Diverse stimuli have been used to tune the properties of polymer monolayers, such as the system temperature, the pH, the ionic strength and the nature of the solvent [153] . In particular, the reaction to the last of these stimuli, i.e. the nature of the solvent, has been extensively explored both theoretically [154] and experimentally [140, 141, 143, 146] , concluding that it significantly affects the structure and the properties of these monolayers. Nevertheless, to the extent of our knowledge, the effect of the water molecules in the properties of these systems, and its variation upon changes in the water concentration has not been explained in detail. Here, we address this void in the literature by analyzing the structural variation with the number of water molecules of a ssDNA monolayer with atomistic MD simulations.

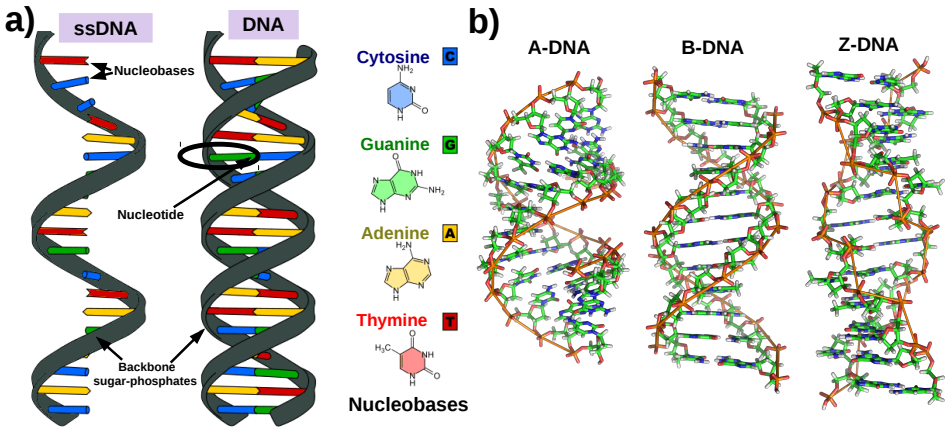



**Figure 3.1:** ***Structural characteristics of the ssDNA and DNA molecules.*** ***a)****Schematic representation of the ssDNA and DNA molecules. The four different nucleobases present in these two molecules are also here shown.* ***b)****Conformation of the A-DNA, B-DNA and Z-DNA molecules.*

### 3.1.2 ssDNA and DNA molecules

The DNA is the nucleic acid which encodes the genetic instructions used in the development and functioning of most living organisms. Additionally, it is the responsible from the hereditary transmission of this information. Its structure consists on two nucleotide strands coiled around each other forming a double helix structure, see Fig. 3.1a. It exists in many possible conformations, that include A-DNA, B-DNA and Z-DNA conformations, see Fig. 3.1b. However, only the last ones (B-DNA and Z-DNA) have been directly observed in functional living organisms [155]. Each conformation can be identified from its diameter (25.5Å, 23.7Å and 18.4Å, respectively), the number of nucleobases in a double-helix turn (11, 10 and 12, respectively), and the distance between two consecutive nucleobases (2.3Å, 3.4Å and 5.3Å).

The term *ssDNA molecule* refers to one of the two nucleotide strands that conform the DNA molecule, see Fig. 3.1a. A nucleotide is the structural component of both ssDNA and DNA and is composed by a monosaccharide sugar (deoxyribose), a phosphate group and a nitrogen-containing nucleobase. This last component, i.e. the nucleobase, is the only element which can change from

one nucleotide to another among four different possibilities: an adenine (A), a thymine (T), a guanine (G) and a cytosine (C) (see Fig. 3.1a). Hence, the ssDNA/DNA sequence can be defined naming only the nucleobase sequence. It is worth mentioning that the two nucleotide strands which conform the DNA molecule must be complementary, i.e. thymine only joins to adenine (and vice versa) and cytosine only joins to guanine (and vice versa).

### 3.1.3 Tuning mechanical response of ssDNA SAMs: Experimental results

ssDNA monolayers have demonstrated a great potential in the biosensing and biomedical fields [156] . In recent years, many biosensing devices have been developed based on the mechanical response of these ssDNA monolayers and its variation when the ssDNA strands join their complementary sequence [130, 157–159]. An example of such kind of devices was built by Mertens et al. in the work reported in ref. 130. In this work, they analyzed the effect of the hydration conditions on the surface stress ($\sigma$) of a gold-coated silicon micro-cantilever functionalized with a densely-packed SAM of ssDNAs [130]. They were able to control the hydration of this system by monitoring its relative humidity (r.h.). Their results showed that its mechanical response went through three well-differentiated stages when the r.h. increases [130], see Fig. 3.2a. From 0 to 10% $r.h.$ (stage I), its surface stress increased up to $\sigma \sim 50$ mN m$^{-1}$, i.e. the cantilever deflected upwards; From 10 to 65% $r.h.$ (stage II), its surface stress decreased more than 100 mN m$^{-1}$, leading to a downward deflection of the cantilever ($\sigma < 0$); Finally (stage III), i.e from 65 to 100% $r.h.$, its surface stress barely changed. This response with the hydration was substantially modified by exposing the ssDNA strands about $\sim$ 1h to their complementary sequence [130], thus indicating that the nanomechanical response of this system with the hydration is sensible to the DNA hybridization process. More concretely, they did not detect an increase of the surface stress of the cantilever at low hydration levels when the ssDNA strands were exposed to their complementary sequence [130], see Fig. 3.2b. Their most interesting result was extracted by exposing the ssDNA strands functionalized in the cantilever to a single-base

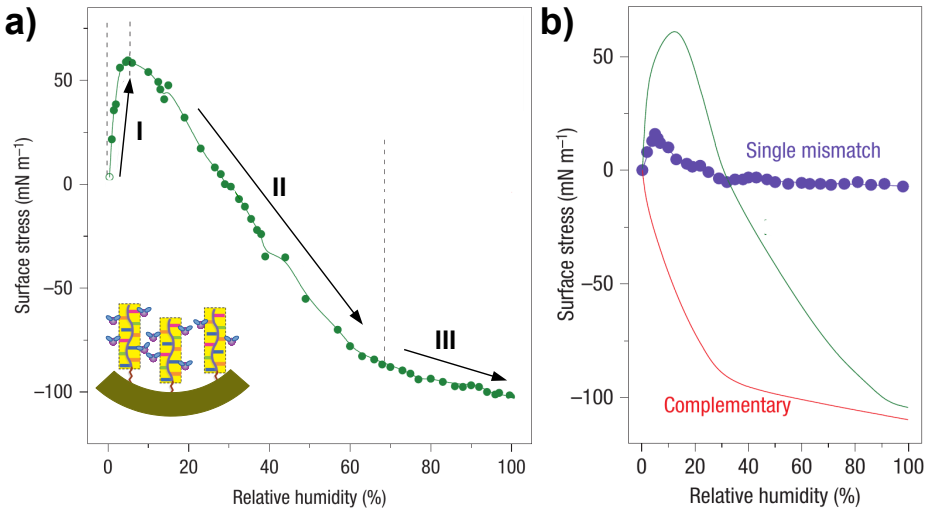



**Figure 3.2:** ***Nanomechanical response with the hydration of a cantilever functionalized with a densely packed SAM of ssDNAs experimentally measured in ref. 130. a)*** *Variation of the surface stress of the cantilever as the r.h. increases. The three well-differentiated stages are here indicated. A schematic representation of the system is shown in the inset.* ***b)*** *Changes observed in the nanomechanical response of this cantilever when the the ssDNA strands are exposed to their complementary sequence (red) or to a sequence with a single-base mismatch (purple). These figures have been reproduced from ref. 130 by permission of Springer.*

mismatched ssDNA sequence [130]. In this case, the nanomechanical response with the hydration was more similar to the one obtained for the ssDNA strands alone than the one detected when they were exposed to their complementary sequence, as the surface stress of the cantilever increased for low r.h. values, see Fig. 3.2b. This last result highlighted the possibility of detecting DNA mutations with an unprecedented sensitivity (single DNA-base) via the nanomechanical response of this biosystem with the hydration.

However, the physical mechanisms behind the mechanical response with the hydration of this SAM-cantilever system are not yet well understood, which is the first step towards its potential use as a biosensing device. Moreover, it has been evidenced that its hydration-driven response strongly depends on its structural factors, such as the ssDNA grafting density [160] or the length of the ssDNA

sequence [161], which complicates even more the task of fully controlling this potential biosensor. It seems that the water-ssDNA hydration forces drive the nanomechanical changes of the whole system, as similar results were observed with an uncharged synthetic mimic of DNA (the PNA molecule) [130], which indicates that the electrostatic interactions between different ssDNAs could not play an essential role in the mechanical response. Still, many important questions about the functioning of this system continue without a clear answer, such as how the water-ssDNA interactions could change as the r.h increases or how that variation could affect the mechanical response of the whole nano-system. That information is difficult to be gathered only with the experiments, which highlights the necessity of analyzing the hydration influence in the mechanical response of a SAM of ssDNAs with theoretical methods for a complete understanding of this problem.

### 3.1.4 Overview

In the last decade, few theoretical works [162, 163] have been developed to explain the origin of the unprecedented hydration-driven mechanical variation measured in ref. 130 (see previous subsection). In both works, they used a mean-field model to understand the influence of water in the structure, free energy and mechanical response of a ssDNA monolayer. The main difference between both works is that the formation of ssDNA-water h-bonds and water-water h-bonds is only considered in ref. 163. Although both were able to capture the experimental non-monotonic response of this system with the hydration [162, 163], the effect of the water molecules on the mechanical response of a SAM of ssDNAs is not yet well understood. In that goal, the use of atomistic MD simulations can suppose a major breakthrough as they are able to accurately describe simultaneously the atomic water-ssDNA interaction and the SAM collective movements promoted by that interaction. Here, we use this method to analyze the hydration effect on the structure and mechanical properties of a ssDNA assembly. To that aim, we have build up a rectangular SAM conformed by 200 ssDNA strands; we have restrained its movement to a substrate whose position is not fixed; and we have analyzed how its curvature changes with the number of water molecules

of the system by performing 180 ns-long MD simulations at different hydration levels. The substrate chosen for performing these simulations was a one-layer graphene, as it is more flexible than the one used in the experiments [130]. Thus, the hydration effects induced on the SAM mechanical response will be amplified, which will allow us to visualize in a much smaller system (nano-scale) the same mechanical effects observed in the experiments (micro-scale). Additionally, we study two ssDNA monolayers with a different grafting density value each, i.e a different number of ssDNA molecules per unit area. In this chapter, we present all the results obtained in this work. The chapter is organized as follows. In sec. 3.2, we expose the details of the MD simulations used for studying the mechanical response of our *SAM-graphene* system at different hydrations, as well as the energetic and structural magnitudes analyzed for understanding the hydration effect. In sec. 3.3, we show that our MD simulations predict a huge variation of the ssDNAs SAM mechanical response with the hydration. Our results agrees with the experiments [130], which validates the force fields used in this work for describing the water-SAM interaction. In sec. 3.4, we analyze all the energetic components that could drive the SAM mechanical response variation with the hydration observed in sec. 3.3, concluding that the water-ssDNA interaction and the water-water interaction are the most relevant ones. We distinguish three hydration ranges regarding the importance of these two interactions at each hydration level. In sec. 3.5, a detailed analysis of the ssDNAs SAM structure at each hydration range is presented. The results obtained from that analysis indicate that by increasing the number of water molecules, the structure of the SAM sharply changes, which influences its curvature. More concretely, we unveil that the hydration-induced changes on the height of the ssDNA strands strongly affect the mechanical response of the whole system. Finally, in sec. 3.6, we analyze the influence of the SAM grafting density on its mechanical response dependence with the hydration. The detailed structural and energetic analysis here presented not only helps to understand the experimental changes observed in the mechanical response of ssDNA monolayers during hydration [130, 160, 161], but also how water could affect to analogous systems, such as polymer brushes.

## 3.2 Methods

### 3.2.1 System preparation and Force Fields

For this work, we have prepared two SAMs constituted by 200 ssDNA strands distributed with a quadrangular packing configuration in a rectangle of 20 x 10 molecules, see Fig. 3.3a. We have chosen a SAM rectangular form to mimic the experiments [130, 160], where the ssDNA strands are immobilized in rectangular micro-cantilevers. The difference between both SAMs is the value of the ssDNAs grafting density ($\rho$), which is $\rho = 4.0\ x\ 10^{13}$ molecules$\cdot\ cm^{-2}$ for the first and $\rho = 3.0\ x\ 10^{13}$ molecules$\cdot\ cm^{-2}$ for the second. These two values of the SAM grafting density are the ones used in the experiments reported in refs. 130, 160 for analyzing the hydration-induced stress variation on ssDNA/DNA monolayers. We have controlled the $\rho$ value by fixing the distance between the bottom-most positions (C5' atom of last nucleotide) of adjacent strands, which is $d_b \sim 15.8(18)$ Å when $\rho = 4.0(3.0)\ \times\ 10^{13}\ cm^{-2}$ (see Fig. 3.3b). Once the SAMs have been prepared, we have centered each of them on top of a one-layer graphene. The dimension of that graphene layer was 30×15.6/35×17.5 nm$^2$ for the SAM with $\rho = 4.0/3.0\ x\ 10^{13}\ cm^{-2}$. The initial distance between the bottommost atom of all the ssDNA strands and the graphene layer was about 3.5Å.

The ssDNA molecule sequence used is the same as in the experimental results reported in refs. 130, 160, i.e. $5' - CTACCTTTTTTTTCTG - 3'$. The initial structure was generated using NAB software [36], which allowed us to reproduce a double helix with the canonical A-form and the desired sequence. Then, we removed the complementary strand, and stretched the ssDNA molecule considering that the diameter of the ssDNA A-form is too large (23Å) [164] to place that strand in such densely-packed SAMs. The final configuration of the strand is about 64Å high and 15.5Å thick (diameter). Given that the phosphate groups in the backbone of the ssDNA are charged, to achieve charge neutrality 15 sodium counter-ions were added per each ssDNA molecule.

The parmbsc0 [165] modification of the Cornell ff99 force field [166] was used for the ssDNA molecules. This force field was known to properly describe

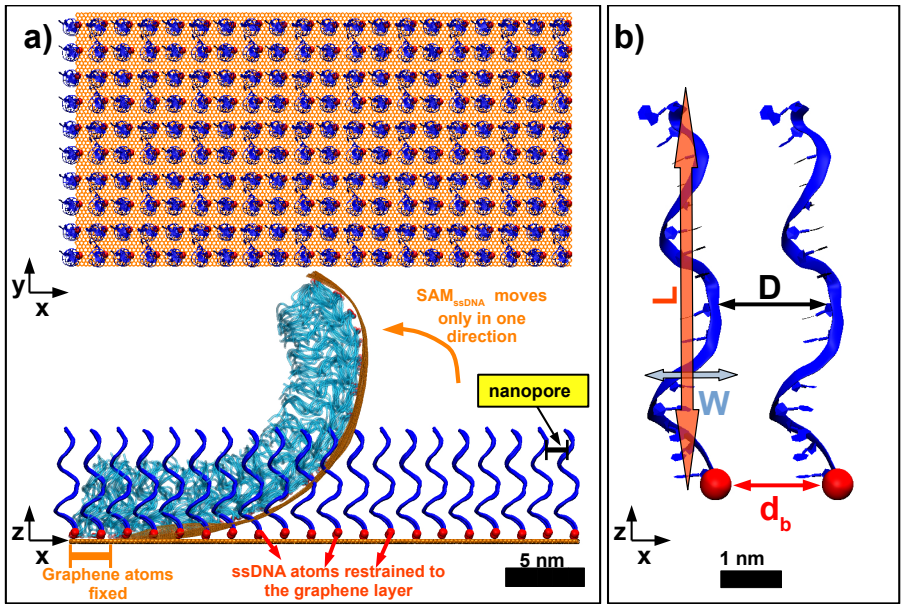



**Figure 3.3:** ***Characterization of the SAM-graphene system used.*** ***a)*** *Top and side views of the system initial configuration. In the side view, we highlight the two main characteristics of our SMD protocol: (1) The graphene atoms whose position is fixed (orange) to force the graphene layer to deflect in only one direction; (2) The bottommost atom of each ssDNA strand (red) whose position is restrained to the graphene layer.* ***b)*** *Definition of the ssDNA length ($L$), the ssDNA width ($W$) and the distance between adjacent ssDNAs ($D$)*

DNA's hydration [167, 168], mechanical, [39] and frictional [169] properties, as well as its behavior in DNA self-assembled monolayers [161]. The graphene atoms were modeled using the OPLS aromatic carbon force field included in the AMBER generalized force field [110]. This force field properly describes graphene mechanical and hydration properties [111, 170] as well as its interaction with biological systems [109, 112] such as DNA/ssDNA [171–173]. The sodium counterions were described using the Joung/Cheatham parameters [174]. The water molecules were described using the Explicit TIP3P water model [114]. This water model suitably described the DNA hydration pattern [168] and also graphene's wetting properties [88].

### 3.2.2 MD parameters and Simulation protocol

We used the AMBER software suite [35] with NVIDIA GPU acceleration [118, 119]. For the sake of numerical efficiency and accuracy, PBC were used with a rectangular box that extends more than $\sim$ 70Å above/beyond the whole system in the x and z directions. To restrict the movement of the system in the y direction, the size of the box in that direction is equal to the system size, i.e. 158.1(180)Å for the SAM with $\rho = 4.0(3.0) \times 10^{13}\ cm^{-2}$. To account for long–range electrostatic interactions, PME [120, 121] was used with a real–space cutoff of 10 Å. vdW interactions were also truncated at the real–space cutoff. We used an integration time step of 1 fs and coordinates were saved every 1000 steps. A constant temperature of 300 K was ensured in all the simulations by means of a Langevin thermostat [48].

Our simulation protocol is composed by 3 main stages. In the first stage we performed an energy minimization to prevent steric clashes, using a combination of steepest descent and conjugate gradient methods. During this process, we kept restraints at all the ssDNA atoms and the graphene monolayer. In the second stage, we heated up the system from 0 to 300 K using a 1 ns long NVT simulation while restraining again the position of all the ssDNA and graphene atoms. Finally, we performed a 180 ns long SMD simulation to analyze the system mechanical behavior. That simulation is based on two main characteristics. First, we kept fix the position of the left-most graphene atoms to force the graphene layer to deflect only in its longer direction, i.e. mimicking the functioning of a micro-cantilever, see Fig. 3.3a. Second, we restrained the distance between the bottom-most atom of all the ssDNA molecules ($C5'$ of the last nucleotide) and its nearer graphene atom to $\sim$ 3.5Å via an harmonic restrain ($k = 100\ kcal/mol$), see Fig. 3.3a. That restrain was maintained during the whole dynamic, allowing that the distance between the ssDNAs bottom-part and the graphene layer hardly changes during the simulation. Thus, we do not restrict the dynamic of the graphene substrate but the ssDNA molecules are always fixed to it, see Fig. 3.3a. We performed that simulation protocol for the two SAMs here studied ($\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ and $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$), and using eleven different hydration levels (H), which are defined accordingly to

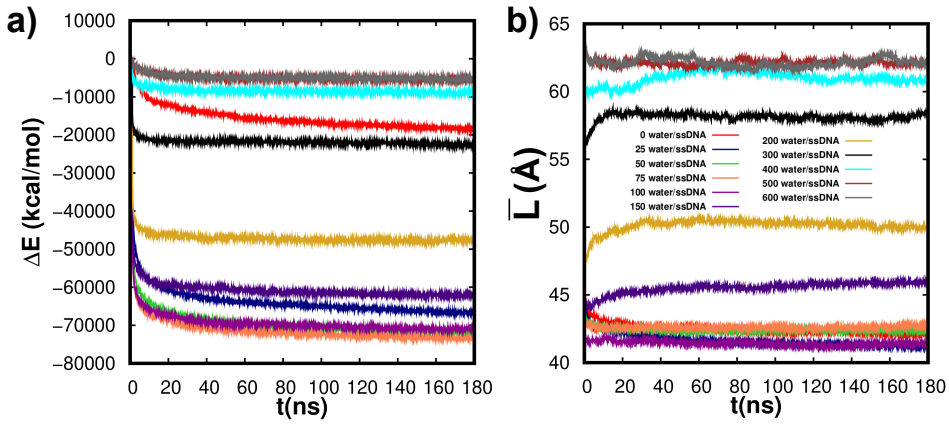



**Figure 3.4:** ***Convergence of the energetic and structural parameters of our system during 180 ns of simulation. a)*** *Time evolution of the total energy change for all the hydration levels studied (0 to 600 water/ssDNA) and using the SAM with* $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$. *The average of the total energy in the first 1 ns has been taken as our energy reference.* ***b)*** *Time evolution of the spatial-average of the ssDNAs length (*$L$*, see sec. 3.2.3) for all the hydration levels here considered (0 to 600 water/ssDNA).*

the number of water per single–stranded DNA (water/ssDNA) that we include in the system: 0, 25, 50, 75, 100, 150, 200, 300, 400, 500 and 600. For all these hydration values, we found that at the end of the simulation (t=180 ns) both the total energy and particularly the structural parameters of the system have fully converged, see Fig. 3.4.

### 3.2.3 Structural and energetic magnitudes analyzed

To estimate the variation of the *SAM-graphene* mechanical response with the hydration, we have evaluated the curvature of this system for all the hydration cases here analyzed. That magnitude has been computed following two steps, see Fig. 3.5. First, we have extracted the x-z profile of the graphene layer at the end of the simulation (t=180 ns). This x-z profile is extracted by averaging the z coordinate value of all graphene atoms with the same x coordinate, see Figs. 3.5a-b. Second, we have fitted that profile to a second-order polynomial, see Fig. 3.5c. Then, the curvature value corresponds to the quadratic term coefficient of that second-order polynomial. Following this protocol for all the

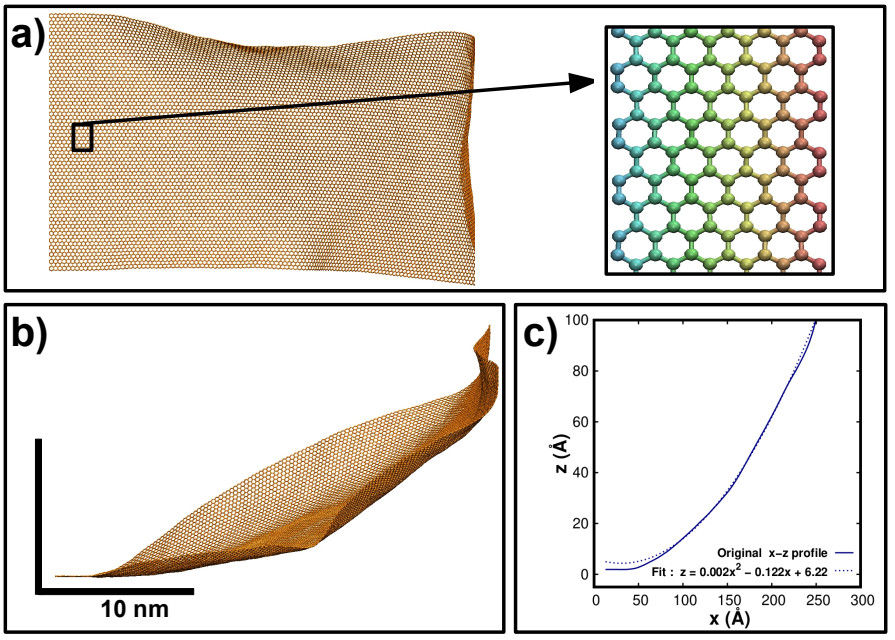



**Figure 3.5:** ***Method used for calculating the SAM-graphene curvature.*** ***a)*** *Top and* ***b)*** *side views of the graphene layer at the end of the MD simulations for $H = 50$. To extract its x-z profile, we average the z coordinate value of all graphene atoms with the same x coordinate. In the inset shown in* ***a)****, we represent some of that graphene atoms colored accordingly to the value of its x coordinate.* ***c)*** *The x-z profile extracted with that graphene coordinates is here shown (continuous blue line). The curvature of that profile is estimated with the quadratic term coefficient of the second-order polynomial which best fits it (dashed blue line).*

hydration cases, we are able to quantify the variation of this system curvature with the hydration.

In order to characterize the structure of the SAM of ssDNAs at each hydration level, we have used the CPPTRAJ tools within the AMBER package. [124]. We have evaluated the four following quantities: the distance between adjacent ssDNAs along the x direction ($D$, see Fig. 3.3b), the SASA [106] of each ssDNA ($SA$), the ssDNAs length ($L$, see Fig. 3.3b), and the ssDNAs width ($W$, see Fig. 3.3b). The length of each ssDNA is estimated using two different methods. First, we have computed the ssDNA end-to-end distance,i.e. $p$. Second, we have calculated all the components of the ssDNAs radius of gyration ($R_g$)

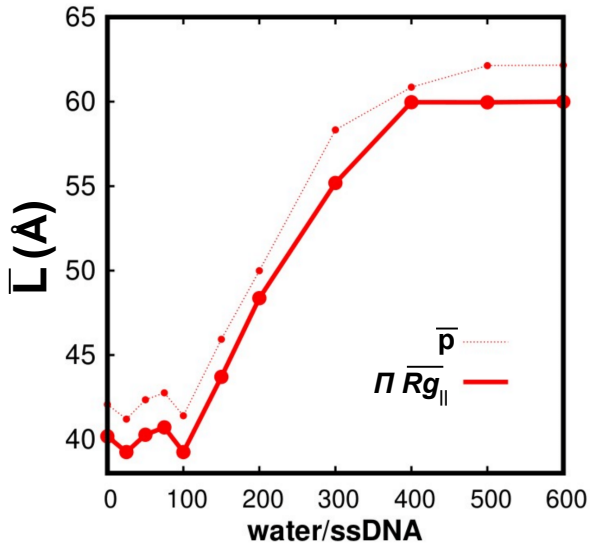



**Figure 3.6:** ***Comparison between the value of*** $p$ ***and the value of*** $\pi \cdot Rg_{||}$ ***for estimating the ssDNA length (*** $L$***).***

tensor[124]. The ssDNA length is proportional to the component of that tensor which is parallel to its axis, i.e. $Rg_{||}$. That is confirmed in Fig. 3.6, where we plot the evolution of the end-to-end distance and $Rg_{||}$ with the hydration when $\rho = 4.0\ x\ 10^{13} cm^{-2}$. As shown in that figure, the ssDNA length can be estimated as:

$$L \sim p \sim \pi \cdot Rg_{||} \tag{3.1}$$

Finally, the ssDNA width is estimated using the components of the $R_g$ tensor that are perpendicular to its axis, $Rg_{\perp}$. Analogously with the $R_g$ length definition shown in eq. 3.1, the relation between $W$ and $Rg_{\perp}$ can be expressed as:

$$W \sim \pi \cdot Rg_{\perp} \sim \pi \cdot \left(\frac{\sqrt{Rg_{\perp,1}^2 + Rg_{\perp,2}^2}}{\sqrt{2}}\right) \tag{3.2}$$

where $Rg_{\perp,1}$ and $Rg_{\perp,2}$ are the two components of the $R_g$ that are perpendicular to the ssDNA axis. Eqs. 3.1 and 3.2, are the ones used to compute the values of $L$ and $W$, respectively, for each ssDNA molecule.

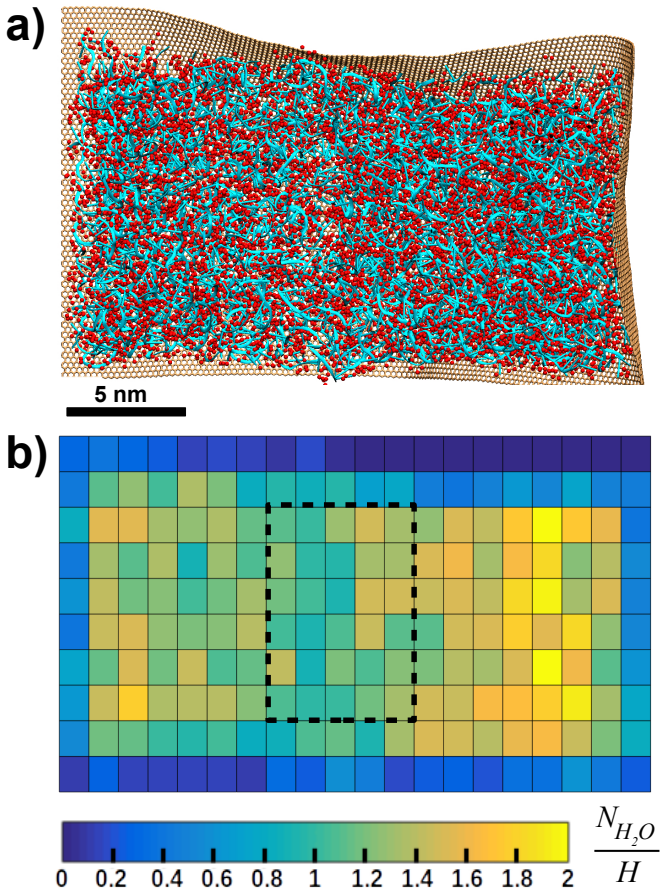



**Figure 3.7:** ***Characterization of the distribution of water molecules along the SAM.*** *The water molecule distribution obtained for $H = 50$ is used as an example.* ***a)*** *The position of the oxygen atoms of all the water molecules (in red) along the SAM at the end of the MD simulation is here shown. The ssDNAs are represented using their backbone configuration coloured in cyan and the graphene atoms are shown in a bond representation colored in orange.* ***b)*** *The histogram of that oxygen atoms distribution along the whole SAM is here shown. The number of water molecules has been normalized to the hydration level (H) of the system. The dashed black rectangle marks the chosen area used to characterize the water distribution at each hydration level.*

We have calculated these four SAM structural magnitudes ($L$, $W$, $D$, $SA$) for all the ssDNA molecules of our SAM. Then, we have computed their spatial-average along the whole SAM and their spatial standard deviation ($\overline{X}$ and $\sigma X$ with $X = L, W, SA, D$). The spatial-average and spatial standard deviation of these structural magnitudes have been evaluated each 20 ps during the last 10 ns of our MD trajectories, i.e. when the system is thermally equilibrated (see Fig. 3.4), and time-averaged in that time slot for studying their evolution with the hydration level (see results in sec. 3.5).

Additionally, we have analyzed the relation between the water molecules and the ssDNAs structure by calculating the number of h-bonds formed between the water molecules and the ssDNA atoms ($N_{hb}$) and by studying the water molecule distribution along the x-y plane of our system (see Fig. 3.3a). In the case of $N_{hb}$, we have computed its total number in the whole system and distinguished between the ones formed with a solely ssDNA atom ($N_{hb}^{s}$) and the ones formed with various ssDNA atoms ($N_{hb}^{b}$). These magnitudes have been evaluated each 20 ps during the last 10 ns of our MD trajectories and time-averaged in that time slot for studying their evolution with the hydration level (see results in sec. 3.5). For estimating the water molecule distribution, we have performed a 20 x 10 histogram of the position of the water oxygen atoms along the whole SAM x-y plane at the end of the simulation, see Fig. 3.7. As shown in that figure, the quantity of water molecules in the SAM limits is systematically lower than in the rest of the SAM. Thus, considering that the water distribution in the limits of the SAM may be significantly distorted, we have defined a region (5x6 histogram cells, see dashed black rectangle in Fig 3.7) placed in the middle of the SAM and we have used that region for comparing the SAM water distribution at different hydration levels.

Finally, we have performed a detailed energy decomposition analysis to unveil the origin of the main driving mechanisms of the *SAM-graphene* collective movement. Considering that this system is composed by three sub-systems (water, ssDNAs and graphene substrate), its total energy can be decomposed in the following six energy terms:

$$E^T = E^G + E^D + E^W + E^{W-D} + E^{G-D} + E^{W-G} \tag{3.3}$$

where $E^G$, $E^D$ and $E^W$ stand for the internal energies of the graphene, the ssDNA strands and the water molecules, respectively; $E^{W-D}$ is the water-ssDNAs interaction energy, $E^{G-D}$ is defined as the graphene-ssDNAs interaction energy and $E^{W-G}$ is the water-graphene interaction energy. These six energy terms were computed as a function of time via an energy decomposition on each recorded frame of our MD trajectories.

The three internal energies are a combination of bonded, electrostatic and vdW interactions. When we refer here to vdW interaction, we mean the interaction described by the Lennard-Jones potential, which includes both the vdW attractive term ($\propto -r^{-6}$) and the Pauli repulsive effects ($\propto r^{-12}$), see sec. 1.2.1. In the case of the three interaction energies, i.e. $E^{W-D}$, $E^{W-G}$ and $E^{G-D}$, they are purely non-bonded interactions, i.e. they only have electrostatic and vdW energy contributions. The water-ssDNAs interaction energy was defined as:

$$E^{W-D}(t) = E^{W+D}(t) - E^W(t) - E^D(t) \tag{3.4}$$

where $E^{W+D}$ stands for the energy of the combined water+ssDNAs system. The other two interaction energies, i.e. $E^{W-G}$ and $E^{G-D}$ were defined analogously.

## 3.3 Mechanical response of the *SAM-graphene* biosystem with the hydration level

In Fig. 3.8b-l, we compare the side view of the final MD configurations obtained after 180 ns of simulation for all the hydration levels here studied, i.e. from 0 to 600 water/ssDNA. From this figure, we observe that the *SAM-graphene* system bends upwards regardless the hydration level. However, its deflection varies more than 25 % between $H = 0$ and $H = 600$. Additionally, the crumpling degree of the graphene layer at low hydration levels is very different to the one observed when $H = 600$. These results confirm that, as was previously observed[130], the mechanical response of this biosystem can be modified by controlling its hydration. To quantify this mechanical change, we compute the curvature of the graphene layer ($\chi$) as explained in sec. 3.2.3 for all the hydration levels here studied, see Fig. 3.8m. At low hydration values, i.e. from $H = 0$ to

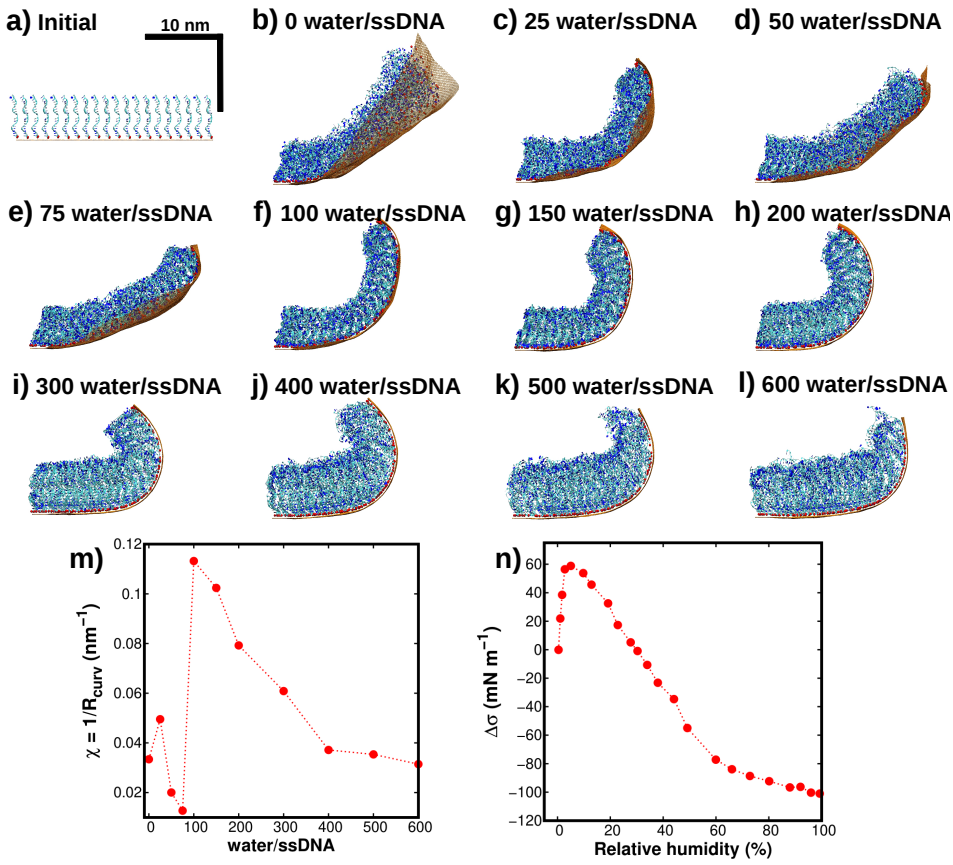



**Figure 3.8:** ***Mechanical response of the SAM-graphene biosystem with the hydration level.*** ***a)*** *Initial configuration of the SAM-graphene system. The graphene atoms are shown in a bond representation colored in orange. The ssDNAs are represented with a ribbon diagram in blue color. The terminal carbons of each ssDNA used to restrain them to the graphene layer are represented in red.* ***b)-l)*** *Side view of the SAM-graphene configurations obtained after 180 ns of simulation for all the hydration levels. The water molecules are not represented for a better visualization of the structure.* ***m)*** *Curvature of the graphene layer at the end of the simulation, $\chi = 1/R_{curv}$, as a function of the hydration level. This magnitude has been computed by fitting the graphene x-z profile to a second-order polynomial (see sec. 3.2.3).* ***n)****Experimental results obtained by J. Mertens et al.[130] .*

$H$ = 75, $\chi$ takes values near $\chi_{not-hydrated} \sim 0.033\ nm^{-1}$. This indicates that the *SAM-graphene* mechanical response does not importantly change in this regime. However, for $H$ = 100, $\chi$ raises up to $3.5 \cdot \chi_{not-hydrated}$, reaching the largest curvature value in the whole hydration range. That result means that the deflection of the *SAM-graphene* system sharply increases from $H$ = 75 to $H$ = 100. Then, for larger hydration levels, $\chi$ gradually decreases until it reaches a saturation value, $\chi_{saturation} \sim \chi_{not-hydrated}$, when $H \sim$ 400, see Fig. 3.8m.

Comparing these results with the experiments reported by Mertens et al. [130], we find that our $\chi$ evolution when $H > 90$ is strongly similar to the cantilever surface stress variation ($\Delta\sigma$) with the relative humidity (r.h.) observed in the experiments, see Fig. 3.8n. This result validates our MD simulations as the surface stress and the curvature of a cantilever must be proportional according to the Stoney's equation[130, 175, 176]. However, the curvature evolution obtained for $H < 90$ in our MD simulations is not observed in the experiments as shown in Fig. 3.8n. This result is expected considering that we use a different method to control the system hydration level, i.e. we change the number of water molecules while they monitor the relative humidity of the system. As was previously reported [177, 178], some water molecules remain bonded to the DNA molecules even at 0% r.h. Therefore, it is impossible to analyze the cantilever mechanical response when the SAM of ssDNAs is completely dehydrated by monitoring the relative humidity. Considering the similarity with the experiments [130] of our MD results above $H$ = 90, see Figs. 3.8m and 3.8n, we conclude that the ssDNAs have $\sim$ 5.6 water molecules per nucleotide firmly bonded to them at 0% r.h. To the best of our knowledge, this is the first estimation of the number of water molecules bonded at 0% r.h to a ssDNA strand. This number is consistent with the 3 water molecules per nucleotide previously reported for a DNA strand [177, 178] considering that in a ssDNA the 2-3 hydrogen bond positions for base pairing are free while they are not in a DNA molecule. Therefore, considering that our MD simulations properly reproduce previous experimental results [130], we trust in the reliability of the huge hydration-driven mechanical change of the *SAM-graphene* system shown

in these simulations.

## 3.4 Analysis of the driving forces of the system during hydration

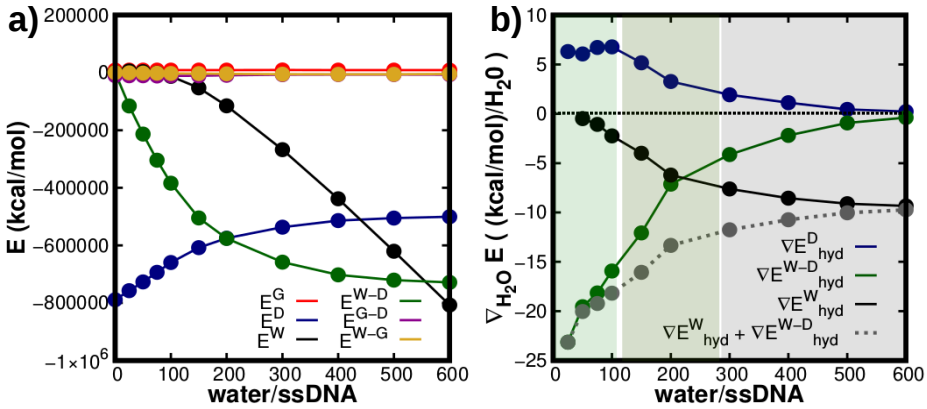



**Figure 3.9:** ***Analysis of the physical origin of the driving forces that cause the SAM-graphene upward deflection. a)****Evolution during the hydration process of the 6 energy components of the SAM-graphene system: graphene internal energy ($E^G$, in red), ssDNAs energy ($E^D = E^{ssDNA} + E^{ssDNA-ssDNA}$, in blue), water energy ($E^W = E^{H_2O} + E^{H_2O-H_2O}$, in black), water-ssDNAs interaction energy ($E^{W-D}$, in green), graphene-ssDNAs interaction energy ($E^{G-D}$, in magenta) and water-graphene interaction energy ($E^{W-G}$, in yellow). The value of each energy component at a given hydration level has been computed by averaging its instantaneous value in the last 10 ns of simulation.* ***b)*** *Evolution with the hydration level of the energetic gradient, $\nabla_{H_2O}\ E_i = \frac{E(hyd_i)-E(hyd_{i-1})}{200\cdot(hyd_i-hyd_{i-1})}$. We compute that energy gradient for the three main energy components:$E^D$, $E^W$ and $E^{W-D}$. The color used for each energy component is the same as in* ***a)****. The energetic gradient arisen from the $E^W$ component when the hydration rises from $H = 0$ to $H = 25$ is not included in the figure as $E^W$ can not be computed when the system is not hydrated. We also represent the sum of $\nabla_{H_2O}\ E^{W-D}$ and $\nabla_{H_2O}\ E^W$ in grey color. The three hydration regimes defined accordingly to the main driving interaction are delimited in* ***b)****: hydR-I in light green, hydR-II in dark green and hydR-III in black.*

To trace back the origin of the *SAM-graphene* mechanical changes observed in the previous section, in what follows, we focus on the evolution of the six en-

ergy components which can influence the behavior of the *SAM-graphene* system: the graphene energy ($E^G$), the ssDNA energy ($E^D$), the water energy ($E^W$), the water-ssDNAs interaction energy ($E^{W-D}$), the graphene-ssDNAs interaction energy ($E^{G-D}$) and the water-graphene interaction energy ($E^{W-G}$), see sec. 3.2. In Fig. 3.9a, we show the evolution with the hydration level of these 6 energy components. As shown in that figure, only three energy components significantly change during the hydration process: $E^W$ (black), $E^D$ (blue) and $E^{W-D}$ (green). Therefore, only these three energetic interactions can drive the mechanical changes observed in sec. 3.3. To quantify their contribution to that driving force, we analyze the energy gradient per water molecule arisen between two consecutive hydration levels for $E^W$,$E^D$ and $E^{W-D}$. The expression used for computing that magnitude ($\nabla_{H_2O}\ E_i$) at each hydration level ($H_i$) is:

$$\nabla_{H_2O}\ E_i = \frac{\Delta\ E}{\Delta N_{H_2O}} = \frac{E(H_i) - E(H_{i-1})}{200 \cdot (H_i - H_{i-1})}. \tag{3.5}$$

where $N_{H_2O}$ refers to the total number of water molecules in the system , defined as $N_{H_2O} = 200 \cdot H_i$.

In Fig. 3.9b, we show the obtained evolution of $\nabla_{H_2O}\ E$ for these three energy components, i.e. $\nabla_{H_2O}\ E^D$, $\nabla_{H_2O}\ E^W$ and $\nabla_{H_2O}\ E^{W-D}$. In the evolution of the ssDNA energy gradient ($\nabla_{H_2O}\ E^D$, in blue) we can distinguish two different phases, see Fig. 3.9b. From $H = 0$ to $H = 100$, this magnitude barely changes with the hydration ($\nabla_{H_2O}\ E^D \sim 6.5$ (kcal/mol)/$H_2O$), while from 100 to 600, it decreases. However, during the whole hydration process, $\nabla_{H_2O}\ E^D > 0$, which indicates that the ssDNA energy does not favor the changes of the *SAM-graphene* configuration observed during hydration. Therefore, $E^D$ not only does not drive the *SAM-graphene* movement described in the previous section, but also it entails a energetic cost for that movement. In contrast, the water energy and the water-ssDNA interaction do promote the *SAM-graphene* mechanical changes analyzed in sec. 3.3, as $\nabla_{H_2O}\ E^{W-D} < 0$ and $\nabla_{H_2O}\ E^W < 0$ during the whole hydration process, see Fig. 3.9b. The sum of these two magnitudes, i.e. $\nabla_{H_2O}\ E^{W-D}$ and $\nabla_{H_2O}\ E^W$ is larger than $\nabla_{H_2O}\ E^D$ for all the hydration levels here considered, see Fig. 3.9b, which confirms that the energetic cost arisen from the $E^D$ component does not hinder

the SAM movement promoted by $E^{W-D}$ and $E^{W}$.

Although both the water energy and the water-ssDNA interaction favor the observed movement of the ssDNAs SAM, two striking differences may be noticed when comparing the evolution of their energetic gradient, see Fig. 3.9b. First, from $H = 0$ to $H = 100$, $|\nabla_{H_2O}\, E^{W-D}| > 6{\cdot}|\nabla_{H_2O}\, E^{W}|$, i.e. the water-ssDNA energy gradient is much larger than the water energy one. This result highlights that, at low hydration levels, the main contribution to the total driving force arises from the water-ssDNA interaction. Second, during the hydration process $|\nabla_{H_2O}\ E^{W-D}|$ decreases to $\sim 0$, while $|\nabla_{H_2O}\ E^{W}|$ increases, taking values near 10 (kcal/mol)/$H_2O$ when $H = 600$. This indicates that the origin of the dominant driving force changes from the water-ssDNA interaction to the water energy during the hydration process, as the former vanishes while the latter strengthens. In fact, we can clearly distinguish three hydration regimes (hydR) regarding the energetic interaction which dominates the SAM driving force, see Fig. 3.9b. First, from $H = 0$ to $H = 100$ (hydR-I), $\frac{|\nabla_{H_2O}\ E^{W-D}|}{|\nabla_{H_2O}\ E^{W}|} > 6$, so the water-ssDNA interaction is the main driving force of the system. Second, from $H = 100$ to $H = 300$ (hydR-II), $0.5 > \frac{|\nabla_{H_2O}\ E^{W-D}|}{|\nabla_{H_2O}\ E^{W}|} > 2.5$, so the main driving force arises from the interplay of both the water-ssDNA interaction ($E^{W-D}$) and the water energy ($E^{W}$). Finally, from $H = 300$ to $H = 600$ (hydR-III), $\frac{|\nabla_{H_2O}\ E^{W-D}|}{|\nabla_{H_2O}\ E^{W}|} < 0.5$, so the energetic contribution which drives the SAM movement is mostly the water energy.

The consequences of the change in the physical origin of the force which drives the SAM of ssDNAs collective movement are better understood considering the definition of the two energy components involved, i.e. the water-ssDNA interaction and the water energy. The first energy component describes the energetic interaction between the water molecules and the ssDNA molecules. This is essentially an electrostatic interaction, i.e. it intensifies when the distance between the water atoms and the ssDNAs is smaller. Thus, when it is the driving mechanism of the system (hydR-I regime), the new water molecules added to the system tend to approach to the ssDNA strands and form new water-ssDNA hydrogen bonds (hereafter refereed as *w-D h-bonds*). Regarding the other driving energetic term, i.e. the water energy ($E^{W}$), it is defined as the internal

energy of the solvent. That energy depends essentially of the number of water h-bonds. Then, when $E^W$ starts to influence the SAM behaviour (hydR-II), the new water molecules added to the system not only tend to approach to the ssDNA molecules, but also tend to increase the number of *water h-bonds* of the system. Finally, in the last hydration regime (hydR-III), the main contribution to the driving force arises from the water energy, which means that the new water molecules enter into the system to form more *water h-bonds* and not to approach the ssDNAs. In summary, the tendency of the new water molecules added to the *SAM-graphene* biosystem is the main driving force of the system, and it changes during hydration. That change can be seen as a three step process: first they tend to approach the ssDNA strands, secondly they tend simultaneously to be near the ssDNAs and increase the number of *water h-bonds*, and finally they only enter into the SAM to interact with the other water molecules and form more *water h-bonds*. In what follows, to put these three water tendencies in contact with the *SAM-graphene* mechanical response observed in sec. 3.3, we study in detail the ssDNAs SAM structural changes promoted by each one of them.

## 3.5 SAM structural characterization during the hydration process

In Fig. 3.10a-b, we show the evolution during hydration of the average value along the whole SAM of three important structural properties of the ssDNAs (see sec. 3.2): its length ($\overline{L}$, in red), its width ($\overline{W}$, in blue) and its SASA ($\overline{SA}$, in purple). $\overline{L}$ evolves from $40 - 60$Å, as shown in Fig. 3.10a. This range correspond to a ssDNA length per base of $2.5 - 3.75$Å, which agrees with previously reported values obtained both experimentally and theoretically for this structural magnitude [161, 179, 180]. In the case of the $\overline{W}$ values, $13 - 16$ Å are obtained, similar to the ssDNA width observed in high-resolution AFM images [181]. The last magnitude, i.e. the $\overline{SA}$, quantifies the SASA of the ssDNAs and is schematically represented with a colored line in Fig. 3.10c. As shown in that figure, the SASA of a ssDNA strand decreases as it is more coiled. Then, the evolution of $\overline{SA}$ provides information about how the ssDNAs structure

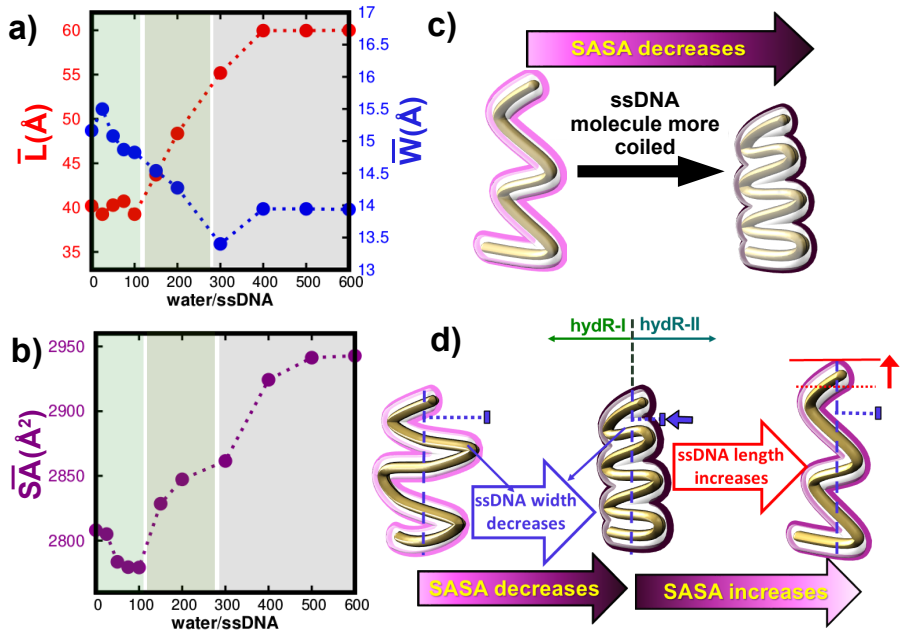



**Figure 3.10:** ***Characterization of the ssDNA structure with the hydration level.*** ***a)-b)****Here we show the evolution with the hydration of the medium value along the whole SAM of:* ***a)*** *the ssDNAs length ($\overline{L}$), the ssDNAs width ($\overline{W}$) and* ***b)*** *the ssDNAs SASA ($\overline{SA}$). The represented value of these magnitudes at each hydration level has been computed by averaging its instantaneous value in the last 10 ns of simulation. The three hydration regimes defined accordingly to the main driving interaction are here delimited: hydR-I in light green, hydR-II in dark green and hydR-III in black. The $\overline{SA}$ estimates how coiled the ssDNA structure is as shown in* ***c)****. The ssDNA SASA in* ***c)*** *is represented with a continuous coloured line.* ***d)****Schematic representation of the ssDNAs structural variation in the hydR-I regime (in light green, left) and in the hydR-II regime (in dark green, right).*

changes during hydration, as it indicates if they are less or more coiled. In the following subsections, we analyze the evolution of these three magnitudes in each hydration regime (hydR-I, hydR-II and hydR-III), to link the three water driving tendencies analyzed in sec. 3.4 with the *SAM-graphene* response observed in sec. 3.3.

### 3.5.1 SAM Structural changes in the hydR-I regime: from 0 to 100 water/ssDNA.

Looking to Fig. 3.10a we observe that $\overline{L} \sim 40$ Å during the whole hydR-I regime , i.e. the medium ssDNA length barely changes with the hydration. Simultaneously, $\overline{W}$ decreases a 4 %, as shown in Fig. 3.10a. Fig. 3.10b shows that $\overline{SA}$ also decreases in the hydR-I regime. Therefore, in this regime, the addition of new water molecules to the system does not affect the ssDNA length, but causes that the ssDNAs become more coiled and narrower over the SAM. This structural ssDNA behaviour is schematically represented in the left panel of Fig. 3.10d.

To understand this change, we must bear in mind that the main driving mechanism in this hydration regime arises from the high water affinity for ssDNA strands (see sec. 3.4). This high water-ssDNA affinity entails that the new water molecules added to the SAM tend to approach to the ssDNA molecules and form new water-ssDNA h-bonds [182, 183]. These *water-ssDNA h-bonds* are not only formed between one water and one ssDNA atom (*single water-ssDNA h-bonds*), but also between one water and various ssDNA atoms (*water-ssDNA bridges*) [184, 185] . This last *water-ssDNA h-bonds* type, i.e the *water-ssDNA bridge* can be formed with ssDNA atoms that belong to the same ssDNA molecule or that form part of different ssDNAs. Therefore, the number of total *water-ssDNA h-bonds* of the system can be defined as:

$$\begin{aligned} N_{hb} &= N_{hb}^{s} + N_{hb}^{b} \\ &= N_{hb}^{s} + N_{hb}^{b,\ intra} + N_{hb}^{b,inter}. \end{aligned} \tag{3.6}$$

where $N_{hb}^{s}$ refers to the number of *single water-ssDNA h-bonds* and $N_{hb}^{b} = N_{hb}^{b,intra} + N_{hb}^{b,inter}$ to the total number of *water-ssDNA bridges*. In Figs. 3.11a and 3.11b, we plot the evolution of both $N_{hb}^{s}$ and $N_{hb}^{b}$, respectively, during the SAM hydration.

From these figures, we observe that both $N_{hb}^{s}$ and $N_{hb}^{b}$ continuously increases from $H = 0$ to $H = 100$, which is consistent with the high water-ssDNA affinity that characterize this hydration range. The $N_{hb}^{s}$ increase is linearly proportional to the water molecules increment at each hydration value of this hydR-I regime

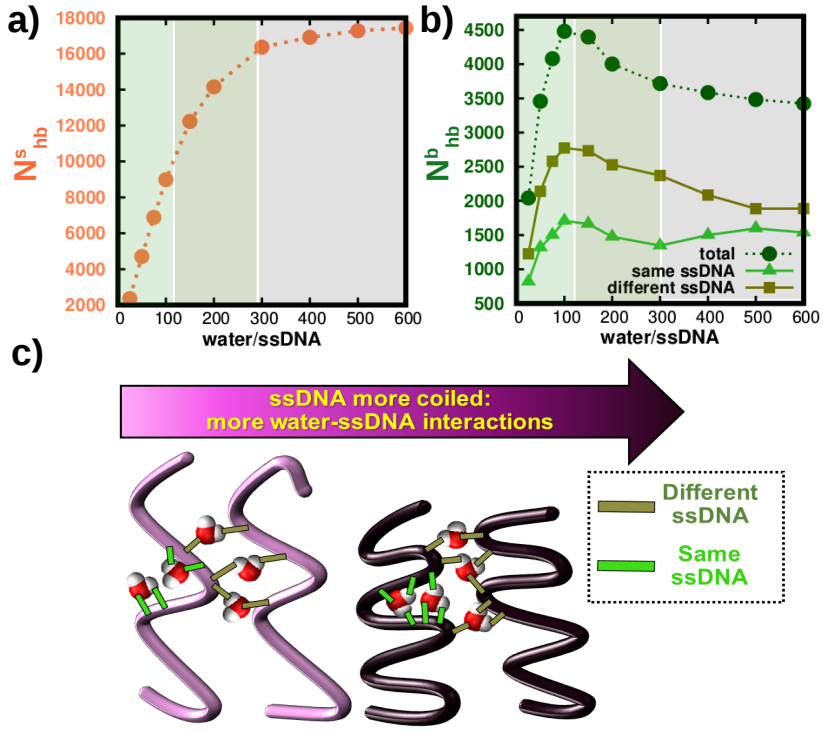



**Figure 3.11:** ***Relation of the ssDNA structure evolution with the total number of water-ssDNA h-bonds.*** ***a)****Evolution during the hydration process of the total number of single water-ssDNA h-bonds, i.e. the number of h-bonds formed between one water and one ssDNA atom.* ***b)****Evolution during the hydration process of the total number of h-bonds formed between a single water molecule and several ssDNA atoms ($N_{hb}^{b}$, dark green line). We distinguish between the number of h-bonds formed with ssDNA atoms that belong to the same ssDNA strand (light green line) and the number of h-bonds formed with atoms that belong to different ssDNAs (olive line). The three hydration regimes defined accordingly to the main driving interaction are here delimited: hydR-I in light green, hydR-II in dark green and hydR-III in black.* ***c)****Schematic representation of the relation between $N_{hb}^{b}$ and how coiled is the ssDNA molecule. The color of the ssDNA structure is as darker as more coiled the ssDNA is.*

($\frac{N_{hb}^{s}}{N_{H_2O}} \sim 0.45$, see Fig. 3.11a). This result confirms that the increase of $N_{hb}^{s}$ is only a direct consequence of the addition of new water molecules to the SAM. Therefore, the system tendency to increase this magnitude is not going to directly affect the ssDNA structure. In contrast, the formation of a new *water-*

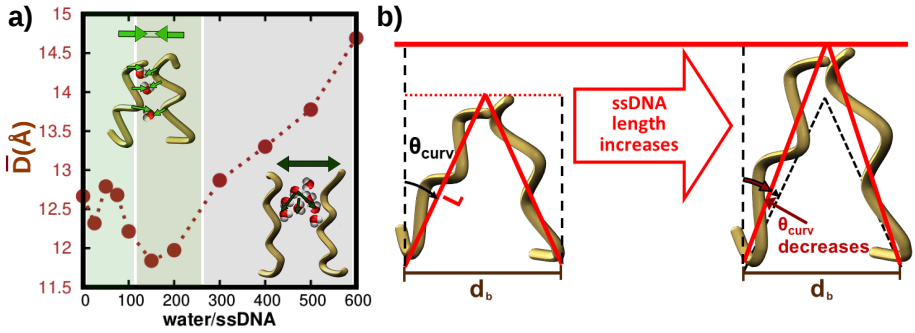



**Figure 3.12:** ***Relation of the SAM structural changes with the variation of its mechanical response a)****Evolution with the hydration of the medium value along the whole SAM of the inter-ssDNA distance ($\overline{D}$). The represented value at each hydration level has been computed by averaging its instantaneous value in the last 10 ns of simulation. The three hydration regimes (hydR-I, hydR-II, hydR-III) defined accordingly to the main driving interaction are here delimited. The relation between $\overline{D}$ and the main driving force of the system in each hydration regime is here depicted with two schematic representations.* ***b)****Schematic representation of the relation between the ssDNAs length and the SAM curvature.*

*ssDNA bridge* requires that some parts of the ssDNAs become closer enough to enable that a solely water molecule form a new h-bond brigde between them. Then, the increase of $N_{hb}^{b}$ (see Fig. 3.11b) does imply changes in the ssDNA structure. In fact, the ssDNA structural changes described in the first paragraph of this subsection are consistent with an increase of $N_{hb}^{b}$. As is schematically represented in Fig. 3.11c, the more coiled the ssDNA structure, the larger the number of both $N_{hb}^{b,intra}$ and $N_{hb}^{b,inter}$ that it can form. Then, from $H = 0$ to $H = 100$, the high water-ssDNA affinity prompts an increase of $N_{hb}^{b}$, which leads to a more coiled configuration of the ssDNAs. This, in turn, results in a reduction of the average ssDNA width along the SAM, as is schematically described in Fig 3.10d.

The increase of $N_{hb}^{b}$ observed in Fig. 3.11b not only explains the ssDNA structural changes, but also the *SAM-graphene* mechanical response variation observed in Fig. 3.8m during the hydR-I regime. As shown in Fig. 3.11b, the increase of $N_{hb}^{b}$ arises from the increase of both $N_{hb}^{b,intra}$ and $N_{hb}^{b,inter}$. The in-

crease of $N_{hb}^{b,intra}$ prompts the stiffening of the ssDNAs structure. This stiffening process of the ssDNA molecule via its interaction with water was previously experimentally observed [184] and causes that it costs more energy to deform its structure. Concerning the increase of $N_{hb}^{b,inter}$, it means that more *water-ssDNA bridges* are formed between adjacent ssDNAs. To that end, the ssDNAs need to be closer enough and they can not separate during this hydration regime. To prove that assumption, we analyze the evolution of the medium value of the inter-adjacent-ssDNAs distance along the whole SAM ($\overline{D}$, see Fig. 3.12a). Looking to that evolution, see Fig. 3.12a, we observe that $\overline{D} \sim 12.5$Å when the system is not hydrated and it barely changes during this first hydration regime. As the bottom position of the ssDNAs is restrained to $d_b \sim 15.8$Å , that result highlights that the adjacent ssDNAs approach together and they do not separate from $H = 0$ to $H = 100$ (see schematic representation in Fig. 3.12a). Therefore, during this first hydration regime the ssDNAs approach to their adjacent ssDNAs up to $\overline{D} \sim 12.5$Å , which is compatible with an upward deflection of the *SAM-graphene* system. What's more, they become stiffer during this range. This indicates that it costs more energy to maintain that close distance between adjacent ssDNAs, prompting an increase of the SAM mechanical response. This explains the results observed in Fig. 3.8m during this hydR-I regime, i.e. an enhancement of the *SAM-graphene* mechanical response.

However, that enhancement of the *SAM-graphene* mechanical response is not a gradual process, see Fig. 3.8m. From $H = 0$ to $H = 75$, the mechanical response barely changes and then it sharply enhances from $H = 75$ to $H = 100$ (see sec. 3.3). In contrast, the ssDNAs structure progressively stiffens during the whole hydR-I regime, as shown in Figs. 3.10b and 3.11b. To understand this difference, we have analyzed the homogeneity grade of the previously analyzed ssDNAs structural quantities along the SAM. In Fig. 3.13a, we show the spatial standard deviation of the ssDNA length ($\sigma L$, top row) and of the inter-adjacent-ssDNAs distance ($\sigma D$, bottom row). As shown in this figure, the homogeneity grade along the whole SAM is completely different from $H = 0$ to $H = 75$ than from $H = 75$ to $H = 100$. In the first range, i.e from $H = 0$ to $H = 75$, both $\sigma L$ and $\sigma D$ are larger or similar to its value when the system is not hydrated.

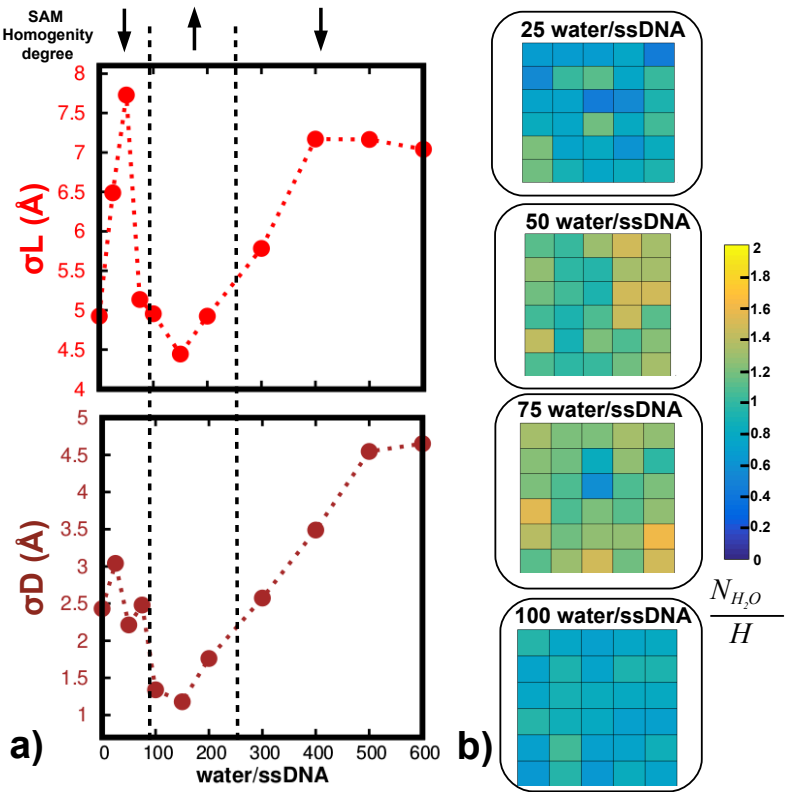



**Figure 3.13:** ***Homogeneity grade of the structure of the SAM. a)****Evolution with the hydration of the spatial standard deviation of two magnitudes which characterize the ssDNAs structure: the ssDNAs length ($\sigma$ $L$, top row) and the inter-ssDNA distance ($\sigma$ $D$, bottom row). The represented value of these magnitudes at each hydration level has been computed by averaging its instantaneous value in the last 10 ns of simulation.* ***b)*** *Water molecules distribution in the center of the SAM (see Fig. 3.7 in sec. 3.2.3). From top to bottom, we show the water molecule distribution when $H = 25$ to $H = 100$.*

This is at odds with the result observed at $H = 100$, when both $\sigma L$ and $\sigma D$ start to decrease. This evolution agrees with the change of the water molecule distribution homogeneity from $H = 0$ to $H = 100$, represented with color maps in Fig. 3.13b (see Fig. 3.7 in sec. 3.2.3). In that figure, we observe that the water molecules distribution along the SAM is rather heterogeneous from $H = 25$ to $H = 75$ but it is remarkably homogeneous at $H = 100$. That compatibility between the water distribution results and the ssDNA structure ones indicates

that it does exist a change of the SAM homogeneity degree in this hydR-I range. From $H = 0$ to $H = 75$, the SAM is quite heterogeneous but from $H = 75$ to $H = 100$ it starts to homogenize. This change in the SAM homogeneity explains why we do not observe a significant enhance of the *SAM-graphene* mechanical response up to $H > 75$ . When the SAM is quite heterogeneous, i.e. from $H = 0$ to $H = 75$, the ssDNAs move in different directions. This causes a non-uniform response of the SAM, which can not prompt an increase of the *SAM-graphene* system mechanical reaction with the hydration even when the ssDNAs are becoming stiffer. On the contrary, from $H = 75$ to $H = 100$, the homogeneity degree of the SAM increases. This means that the ssDNAs start to work cooperatively in the same direction, what leads to the sharp increase of the *SAM-graphene* mechanical response observed in Fig. 3.8m.

To conclude, in this subsection we have been able to explain both the structural and the mechanical behaviour of the SAM in the hydR-I range, considering that in this regime the main driving force arises from the high water-ssDNA affinity. As observed in Figs. 3.11a and 3.11c, that high water-ssDNA affinity causes an increase of both $N^s_{hb}$ and $N^b_{hb}$, which in turn leads to two different ssDNAs structural variations: they become stiffer and move towards their adjacent strands. These factors prompt that, when the SAM homogeneity is high enough (from $H = 75$ to $H = 100$), the mechanical response of the *SAM-graphene* system sharply increases.

### 3.5.2 SAM Structural changes in the hydR-II regime: from 100 to 300 water/ssDNA.

To understand also the *SAM-graphene* mechanical changes happenning during the hydR-II regime, now we focus on the ssDNA structural quantities evolution during this range, i.e. from $H = 100$ to $H = 300$. In Fig. 3.10a, we observe that $\overline{L}$ starts to substantially increase in this hydration regime. It raises from $\sim 40$Å to $\sim 55$Å, which is more than a $35\%$. That increase is accompanied by a $\overline{W}$ decrease, see Fig. 3.10a, and an increment of the $\overline{SA}$, see Fig. 3.10b. This result indicates that in this hydration range, the addition of new water molecules to the system stretches the ssDNA molecules, causing them to become narrower

and less coiled. That structural ssDNAs behavior is schematically represented in the right panel of Fig. 3.10d.

Looking now to both $N_{hb}^{s}$ and $N_{hb}^{b}$ evolution, see Figs. 3.11a and 3.11b respectively, we observe that in the hydR-II range the former continues increasing while the latter starts to decrease. This is consistent with the system driving mechanism in this hydR-II regime, see sec. 3.4. In this hydration range, the new water molecules added to the system have still affinity for the ssDNA strands, which explains the $N_{hb}^{s}$ increase. However, these new water molecules also tend to interact with other water molecules and form new *water h-bonds*, see sec. 3.4. This indicates that in the hydR-II regime the water-ssDNA affinity is not as high as in the previous regime, which explains the $N_{hb}^{b}$ decrease.

This decrease of $N_{hb}^{b}$ is compatible with the ssDNAs structural results discussed in the previous paragraphs. Considering the connection between the $\overline{SA}$ and $N_{hb}^{b}$ explained in Fig. 3.11c, the reduction of the number of *water-ssDNA bridges* in the system allows for the $\overline{SA}$ to increase, i.e. the ssDNAs configuration become less coiled. That in turn allows that the other two ssDNA structural magnitudes ($\overline{L}$ and $\overline{W}$) evolve to increment the number of water h-bonds of the system. As observed in Fig. 3.10a, $\overline{L}$ increases and $\overline{W}$ decreases in this hydR-II range, which is translated into a larger free volume between two adjacent ssDNA molecules. That free volume constitute a nanopore of the SAM (see Fig. 3.3 in sec. 3.2.1) and the larger it is, the more water molecules enter inside it. Therefore, during the hydR-II range the ssDNAs not tend to be more coiled but to stretch, letting more water molecules enter to the SAM nanopores and thus promoting that the number of *water h-bonds* increases.

Finally, we have related this ssDNAs structural variation with the SAM mechanical response observed in this hydR-II regime. As shown in Fig. 3.8m, that mechanical response decreases from $H = 100$ to $H = 300$, which can be explained considering two different SAM structural aspects. First, the inter-adjacent-ssDNA distance ($\overline{D}$) keeps near $\sim$ 12Å during this whole second regime, see Fig. 3.12a. This result means that adjacent ssDNAs preserve the relative position that they have in the first hydration regime, i.e. they tilt towards each other (see sec. 3.5.1 and the schematic representation of Fig. 3.12a).

Second, $\overline{L}$ greatly increases as is shown in Fig. 3.10a. That increase reduces the maximum curvature angle accordingly to the equation (see schematic representation in Fig. 3.12b):

$$\theta_{curv} \propto \frac{d_b}{2 \cdot \overline{L}} \tag{3.7}$$

Therefore, the increase of the length of all the ssDNAs necessarily implies that the curvature variation induced by each of them is smaller in the second regime than in the first (hydR-I). This explains the decrease of the *SAM-graphene* mechanical response shown in our MD simulation results (see Fig. 3.8m).

To sum up, the coexistence of both water-ssDNA and water-water affinities that characterizes this hydR-II range prompts that the ssDNAs become less coiled than in the previous regime. Instead, the ssDNA molecules tend to stretch, allowing that more water molecules enter into the SAM nanopores and thus the number of *water h-bonds* increases. That stretching process of the ssDNAs causes the reduction of the maximum curvature that they can induce to the whole SAM, prompting the reduction of the *SAM-graphene* mechanical response even when the ssDNAs continue moving towards each other (see 3.12a). This result evidences that the reduction of the mechanical response does not arise as a consequence of a water-mediated repulsive interaction between ssDNAs as was previosly suggested [130], but from a strong variation of the ssDNAs conformation. Therefore, the general changes induced in the ssDNAs structure during SAM hydration greatly affect the SAM mechanical response, so they must also be considered to understand that collective SAM behavior.

### 3.5.3 SAM Structural changes in the hydR-III regime: from 300 to 600 water/ssDNA.

We now proceed with the analysis of the ssDNAs structural variation during the hydR-III regime, i.e. from $H = 300$ to $H = 600$. In Figs. 3.10a-b, we observe that the evolution in that hydration interval of the $\overline{SA}$, the $\overline{L}$ and the $\overline{W}$ is very similar. From from $H = 300$ to $H = 400$, they increase about a $5\%$. Then, for $H > 400$, their evolution saturates, i.e. they barely change as the hydration raises. Similarly, $N_{hb}^{b}$ and $N_{hb}^{s}$ vary about a $6\%$ from $H = 300$ to $H = 400$ and

they saturate for higher hydration values, see Figs. 3.11a and 3.11b. In contrast with that trend, the value of $\overline{D}$ does not saturate during the hydR-III regime, see Fig. 3.12a. In fact, its evolution drastically changes. While in previous hydration ranges the $\overline{D}$ value remains around $12$Å, in this hydR-III regime it increases up to $\overline{D} \sim 14.75$Å, revealing that adjacent ssDNAs separate from each other from $H = 300$ to $H = 600$.

To understand these results, we must bear in mind that the main driving mechanism of this hydR-III range is the strong water tendency to increase the number of *water h-bonds* (see sec. 3.4). This characteristic implies that the new water molecules added to the system tend to approach to other water molecules. At the beginning of the hydration range, i.e. from $H = 300$ to $H = 400$, most water molecules are inside the SAM nanopores, see Fig. 3.14. Therefore, the high water-water affinity prompts that the new water molecules enter into the SAM nanopores even when there is not enough space for them. That entrance entails an enlargement of the steric pressure of the system, which can only be reduced by separating the adjacent ssDNA strands (see schematic representation in Fig. 3.12a). This explains the $\overline{D}$ evolution during this regime, i.e. the increment of the inter-ssDNA distance for $H > 300$ (Fig. 3.12a).

The forced entrance of water molecules into the SAM explains the evolution of the number of *water-ssDNA h-bonds* at the beginning of this hydration range, i.e. from $H = 300$ to $H = 400$. As the adjacent ssDNAs have to separate from each other, the number of *water-ssDNA bridges* formed between them significantly decreases. That result is confirmed in Fig. 3.11b, where we observe a remarkable reduction of $N_{hb}^{b,inter}$ from $H = 300$ to $H = 400$. This reduction must be compensated by the increase of other type of *water-ssDNA h-bonds*, as it is more favorable to the system not to lose water-ssDNA interactions (see Fig. 3.9a in sec. 3.4). To that end, the ssDNAs structure varies from $H = 300$ to $H = 400$, see Figs. 3.10a-b. As shown in that figures, both $\overline{L}$ and $\overline{W}$ increases, which increments the number of water-ssDNA interactions on the SAM top and inside the nanopores, respectively. This enable that both $N_{hb}^{s}$ and $N_{hb}^{b,intra}$ increase, see Figs. 3.11a-b, compensating the *water-ssDNA bridges* loss promoted by the ssDNAs separation.

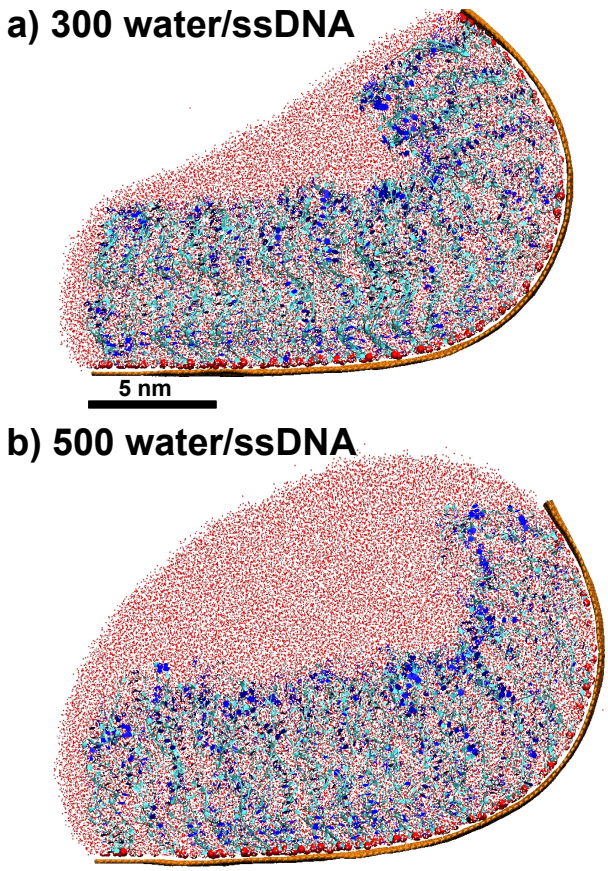



**Figure 3.14:** ***Water distribution variation from $H = 300$ to $H = 600$ (hydR-III range).*** *Side view of the SAM-graphene final configurations obtained after 180 ns of MD simulation using* ***a)*** $H = 300$ *and* ***b)*** $H = 500$. *As shown in this figure, the number of water molecules located outside the SAM greatly increases with the hydration in the hydR-III range. The water molecules are shown in this figure via a point representation in red color. The SAM representation is the same used in Fig. 3.8.*

It is worth to mention that the evolution of the water molecules distribution during the hydR-III range changes, see Fig. 3.14. As shown in that figure, the number of water molecules located on top of the SAM greatly increases for $H > 400$. As a consequence, the formation of new *water h-bonds* with water molecules that are not in the SAM nanopores becomes more favourable as the system hydrates. In other words, the tendency of the new water molecules added

to the system changes from moving inside the nanopores to being placed on top of the SAM. This prompts the reduction of the number of water molecules entering into the SAM nanopores, which directly impacts the SAM structural evolution. First, it progressively diminishes the hydration effect on the number of *water-ssDNA bridges* of the system, see Fig. 3.11b. This decreases the necessity to form more $N_{hb}^{s}$ when $H > 400$, allowing that the ssDNAs structure does not have to change. This is compatible with the results observed in Figs. 3.10a-b for $H > 400$, as the evolution of the three magnitudes which describe the ssDNAs structure saturates. Finally, it is important to mention that the inter-ssDNA distance continues increasing even at $H = 500$. This result highlights that some water molecules continue entering into the nanopores at high hydration levels.

To conclude, we use all the results discussed in this subsection to explain the variation of the *SAM-graphene* mechanical response observed in the hydR-III range, see Fig. 3.8m. From that results, we can extract two main conclusions. First, the increase of the inter-ssDNA distance promoted by the water tendency of entering into the SAM nanopores implies that the ssDNA molecules stop moving towards each other. That phenomenon translates into a reduction of the *SAM-graphene* mechanical response, which agrees with the results depicted in Fig. 3.8m. Second, two stages can be differentiated in the evolution of that *SAM-graphene* response: from $H = 300$ to $H = 400$ it decreases more than a $35\%$, while for $H > 400$ it roughly saturates. That evolution is more similar to the ssDNAs structural variation than to the inter-ssDNA distance one, as only in the former we also distinguish two different stages (see Figs. 3.10a-b). This indicates that the variation of the ssDNAs structure is also influencing the *SAM-graphene* mechanical response. Considering the relation between the length of a ssDNA and the maximum curvature angle that it can induce to the whole system, see sec. 3.5.2 and Fig 3.12b, the cause of these two different stages is revealed. From $H = 300$ to $H = 400$, not only $\overline{D}$ increases but also $\overline{L}$ increases, see Fig. 3.10a. As both changes hinder the SAM mechanical response, the sum of its effects results in a large reduction of the *SAM-graphene* mechanical reaction. However, for $H > 400$ only $\overline{D}$ increases, prompting a

subtle reduction of the SAM mechanical response. This highlights once again the importance of the ssDNAs structural changes during hydration to explain the variation of the whole SAM mechanical response, as it can be only completely understood if these changes are considered.

## 3.6 Effect of the grafting density

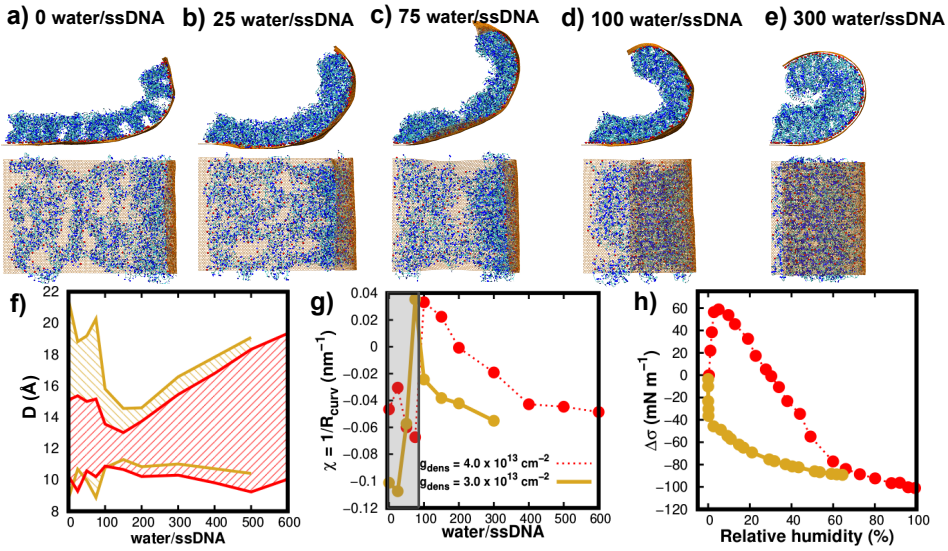



**Figure 3.15:** ***Mechanical response results for the SAM with*** $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ ***obtained after 180 ns of MD simulation. a)-e)*** *Side and top views of the SAM-graphene configurations obtained when hyd = 0, 25, 75, 100, 300 water/ssDNA. The color representation used for the whole system is the same as in Fig. 3.8a-l. The water molecules are not represented for a better visualization of the structure.* ***f)****Evolution with the hydration of the range of inter-ssDNA distance values* $(D)$*, when* $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ *(yellow) and when* $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ *(red).* ***g)****Evolution of the system curvature as a function of the hydration level* ($\chi = 1/R_{curv}$*, yellow continuous line). The result obtained for the SAM with* $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ *is also shown for comparison (red dashed line). Both curves have been rescaled considering that the zero curvature is obtained when* $H = 90$*. The not-accessible hydration range in the experiments[130, 160] , i.e. from* $H = 0$ *to* $H = 90$*, is marked in grey.* ***h)****Experimental results obtained for the variation of the SAM mechanical response with the r.h. when* $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ *[160] (yellow), and when* $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ *[130] (red).*

The previous section reveals that the mechanical response of a densely packed SAM during hydration is closely connected with the structural changes that its ssDNAs suffer during this process. In what follows, we focus on the effect of other SAM structural characteristic which also influences its mechanical response as demonstrated in refs. 130, 157, 160, 186, that is the SAM grafting density. To that aim, we simulate at different hydration levels the dynamic of a *SAM-graphene* system which grafting density is only $\sim$ 1.3 times smaller, i.e. $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$. In Figs. 3.15a-e and 3.16a we show both the side and top views of the final configuration obtained for the eight hydration levels considered for this case, i.e. from $H = 0$ to $H = 300$. From that figures, we observe that the ssDNAs distribution along the SAM at $H = 0$ is more irregular than in the $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ case, see top views in Fig. 3.16a-b. To understand that result, we must consider that the ssDNAs need to approach to each other to keep their helical structure when the system is not hydrated [161] (see sec. 1.1.1). In the previous grafting density case, the ssDNAs are placed near enough [161] to approach their adjacent strands maintaining their regularly arranged initial position in the SAM, i.e. the inter-ssDNA distance is practically the same for all the ssDNAs (see Fig. 3.15f). However, when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$, the distance between the bottom position of the adjacent ssDNAs is larger than in the previous case ($d_b \sim 18$Å $> 15.8$Å). Therefore, the collective movement of all ssDNAs to approach to its adjacent ones is less energetically favorable. This causes that all ssDNAs can not place equidistant from their adjacent ones forming a regularly ordered SAM. Instead, they self-organize in various ssDNA aggregates, which are smaller than the system size, see Fig. 3.15a. As a consequence, the distance between adjacent ssDNAs belonging to distinct aggregates is quite larger than the one corresponding to the same aggregate, leading to an increase of the dispersion of the $D$ value along the SAM, see Fig. 3.15f. This irregularity is reduced when the hydration increases, as the water molecules help to keep the ssDNA strands upright in their regularly arranged initial position [161] .

In Figs. 3.15a-e we also observe that this SAM with $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ bends upwards for all the hydration levels, but its deflection grade strongly changes with the hydration. This indicates that the mechanical response of

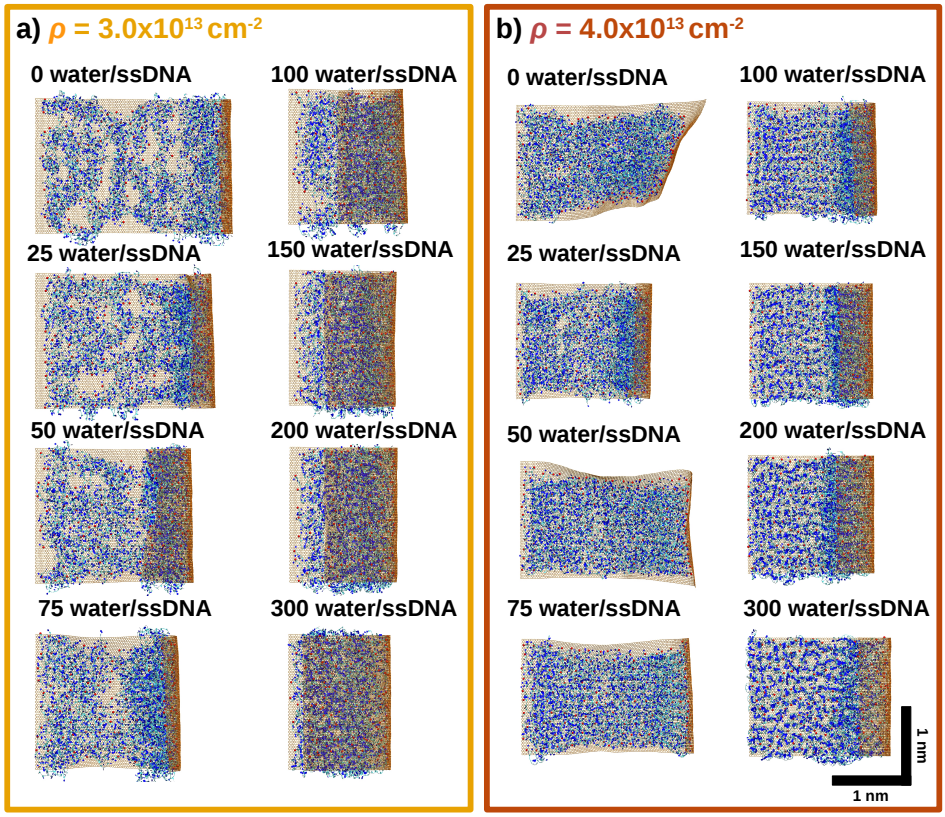



**Figure 3.16:** ***Top views of the SAM-graphene configurations obtained after 180 ns of simulation using a)*** $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ ***and b)*** $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ ***for*** $H$ ***= 0-300.*** *The color representation used for the whole system is the same as in Fig. 3.8a-l. The water molecules are not represented for a better visualization of the structure.*

this system can also be tuned with the hydration. To quantify that mechanical response variation, we use the same methodology as in sec. 3.3, i.e. we estimate the graphene curvature ($\chi$) for each hydration level, see Fig. 3.15f. The $\chi$ evolution obtained can be compared with the experimental results reported by Dominguez et al. in ref. 160. In that work, they analyze the surface stress variation of a SAM with $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ as the r.h. increases, see Fig. 3.15g. As discussed in sec. 3.3, the system curvature evolution when $H > 90$ must be proportional to the experimental surface stress variation for r.h $> 0\ \%$, as

90 water molecules are firmly bonded to each ssDNA strand at 0% r.h. That result is confirmed in Fig. 3.15f, where we observe a $\chi$ reduction when $H > 90$, analogously to the surface stress tendency detected experimentally [160], see Fig. 3.15g. This similarity highlights the relevance of our MD simulation results, as they are compatible with the experiments regardless the SAM grafting density value.

Comparing now this $\chi$ evolution with the one observed when the grafting density is higher ($\rho = 4.0\ x\ 10^{13}\ cm^{-2}$), we observe two similarities, see Fig. 3.15f. For both grafting densities, $\chi$ increases up to a maximum curvature value when the hydration is low, and $\chi$ progressively decreases from $H = 100$ to $H = 300$. Despite of these similarities, two main differences are observed. First, the hydration value that corresponds to the maximum curvature, hereafter refereed as $H_{max}$, is smaller in this new grafting density case. As shown in Fig 3.15f, for this new grafting density $H_{max} = 75$, while $H_{max} = 100$ in the $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ case. Second, in the $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ case we observe a pronounced $\chi$ reduction immediately after the maximum curvature, which is not detected in the higher grafting density case. This difference is also observed when comparing the experimental results obtained for the mechanical response of these two SAM [130, 160], see Fig. 3.15g. What's more, the $H_{max}$ difference explains why an enhancement of the SAM mechanical response with the hydration is only experimentally observed when $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$, see Figs. 3.15f-g. Therefore, our MD simulation results can capture the differences experimentally observed in the mechanical response of SAMs with distinct grafting density values.

To gain further insight into the causes of that mechanical response's differences, we evaluate the effect of the grafting density variation on the competition between the two driving interactions of the system, i.e the water-ssDNA interaction ($E^{W-D}$) and the water-water affinity ($E^{W}$). In Fig. 3.17a, we plot the evolution of the energetic gradient arisen from both interactions during the hydration process, i.e $\nabla_{H_2O}\ E^{W-D}$ and $\nabla_{H_2O}\ E^{W}$, when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$. As shown in that figure, no substantial changes exist in the evolution of that two magnitudes when the SAM grafting density is modified from $4.0\ x\ 10^{13}\ cm^{-2}$ to

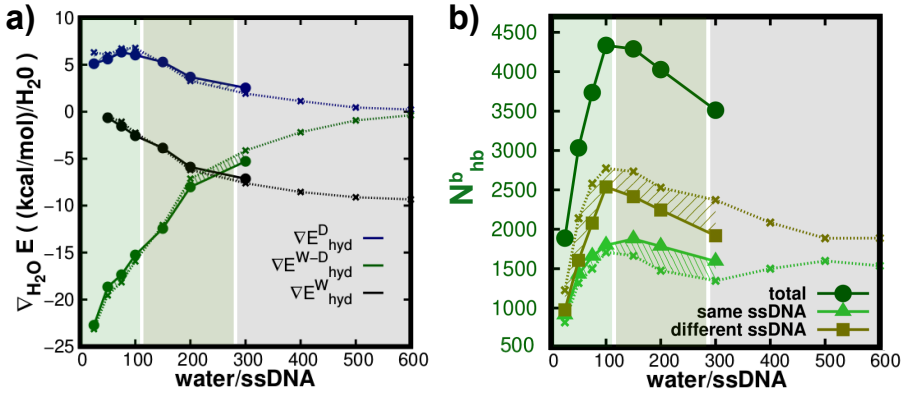



**Figure 3.17:** ***Evolution of a)the energetic gradient ($\nabla_{H_2O}\ E$) arisen from the three main energy components of the system and b)the number of h-bond formed between a single water molecule and several ssDNA atoms ($N^b_{hb}$), for the two grafting densities here studied.*** *For these two magnitudes, we show both the results obtained when $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ (dashed line) and when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ (continuous line). In **b)**, we distinguish between the number of water-ssDNA bridges formed with ssDNA atoms that belong to the same ssDNA strand (light green line) and the number of of water-ssDNA bridges formed with atoms that belong to different ssDNAs (olive line). The three hydration regimes defined accordingly to the main driving interaction are here delimited: hydR-I in light green, hydR-II in dark green and hydR-III in black.*

$3.0\ x\ 10^{13}\ cm^{-2}$ (see Fig. 3.9 in sec. 3.4). This result indicates that the three hydration ranges (hydR-I,hydR-II, hydR-III) delimited for $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ can also be used in this new case, as the evolution of the main driving interactions with the hydration is essentially the same. More importantly, it highlights the necessity of a deeper analysis of the properties of these SAMs to explain the mechanical response differences observed between them.

We continue tracing the origin of the two differences in the mechanical response by analyzing the effect of this grafting density change on the ssDNAs structural evolution. In Figs. 3.18a-b, we plot the evolution of the $\overline{L}$, $\overline{W}$ and $\overline{SA}$ during hydration when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$. We focus on the evolution of these three magnitudes from $H = 0$ to $H = 100$, i.e the hydration range where the two mechanical response differences are detected (hydR-I). In that

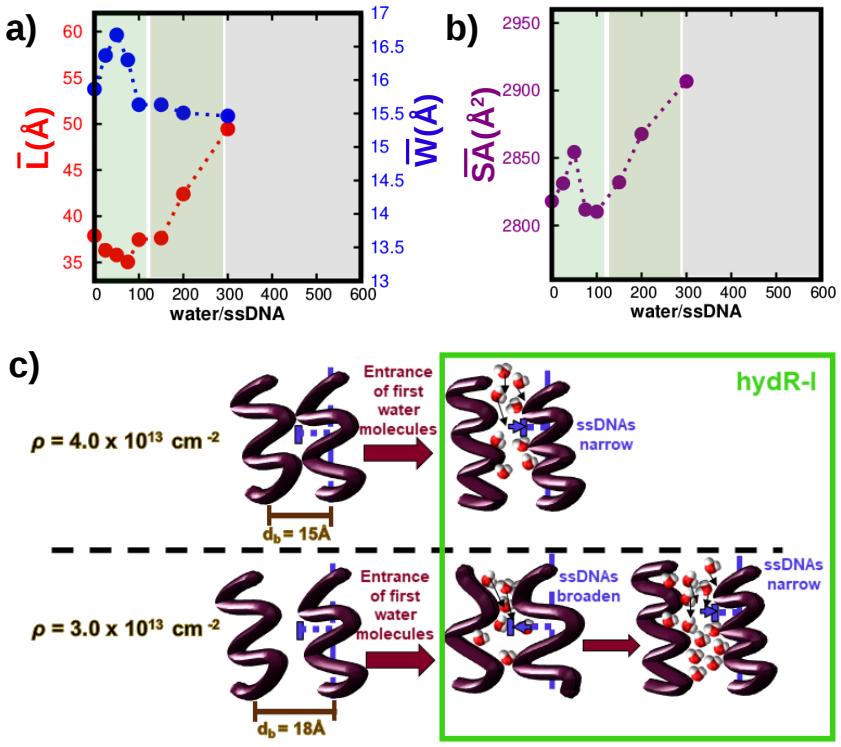



**Figure 3.18:** ***Structural characterization of the ssDNA structure with the hydration level for the*** $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ ***case.*** ***a)-b)****Here we show the evolution with the hydration of the averaged value along the whole SAM of:* ***a)*** *the ssDNAs length* ($\overline{L}$)*, the ssDNAs width* ($\overline{W}$) *and* ***b)*** *the ssDNAs SASA* ($\overline{SA}$)*. These values at each hydration level have been computed by averaging its instantaneous value in the last 10 ns of simulation. The three hydration regimes defined accordingly to the main driving interaction are delimited in both* ***a)*** *and* ***b)****: hydR-I in light green, hydR-II in dark green and hydR-III in black.* ***c)*** *Schematic representation of the structural differences observed between the* $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ *case (top panel) and the* $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ *case (bottom panel) during the hydR-I range.*

hydration range, we observe that both $\overline{W}$ and $\overline{SA}$ increases from $H = 0$ to $H = 50$, at variance with the results obtained when $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$. What's more, the $\overline{L}$ decreases for low hydration values ($H = 0 - 75$) when in the $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ case this magnitude barely changes during the whole hydR-I range (see Fig 3.10a). These differences evidence that the SAM structural

behavior during the hydR-I range depends on the grafting density value, even when the main driving force of the system is in both cases the formation of new *water-ssDNA h-bonds*.

To understand that dependency, the variation of the initial inter-ssDNA distance value when the SAM grafting density is changed must be considered. In the higher density case, the ssDNA strands are placed at $d_b = 15.8$ Å. When the system is not hydrated, $W$ is approximately equal to $d_b$, see Fig 3.10a. This indicates that all the space between two ssDNA strands is essentially filled with the atoms of these two ssDNAs, see top panel of Fig. 3.18c. As the first water molecules are introduced in the system, the ssDNAs structure becomes narrower and more coiled, allowing the entrance of that water molecules into the nanopores and the increase of the *water-ssDNA h-bonds* number. However, $d_b \sim 18$ Å in the $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ case. That value is larger than the ssDNA width when the system is not hydrated, which is $W \sim 15.7$Å, see Fig 3.18a. This means that the space between adjacent ssDNAs at $H = 0$ is not completely filled by their ssDNA atoms, see Figs. 3.15a and 3.18c. Consequently, the ssDNAs do not need to narrow for letting the entrance of the first water molecules into the nanopores. In fact, they tend to broaden, allowing a larger ssDNA-water-ssDNA interaction inside the nanopores and enhancing the number of *water-ssDNA bridges* formed between different ssDNAs, see Fig. 3.17b. This broadening process of the ssDNAs structure entails that they naturally shorten and become less coiled, which explains the obtained evolution at very low hydration levels of the $\overline{L}$ and the $\overline{SA}$, respectively (see Figs. 3.18a-b). Thus, the larger inter-ssDNA distance value which characterizes the SAM with $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ explains the ssDNAs structural changes observed for this case at low hydration levels and why they are not detected when $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$.

The ssDNAs broadening process continues until the number of water molecules is large enough to fill the volume of the nanopores, i.e. when $H \sim 50 - 75$. From that point, the high water-ssDNA affinity which characterizes the hydR-I range is translated into other structural tendency. As shown in Fig 3.18a, the $\overline{W}$ and the $\overline{SA}$ decreases from $H = 50$ to $H = 100$, similarly to the struc-

tural evolution obtained in the $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ case during the whole hydR-I range, see Fig 3.10a. This result indicates that once the volume of the nanopores is filled, the ssDNAs continue enhancing the water-ssDNA interaction using the same method than in the previous grafting density case, i.e. becoming narrower and more coiled and thus promoting the formation of more *water-ssDNA bridges*, see Fig. 3.17b. This new ssDNAs structural tendency, is compatible with the increase of $\overline{L}$ observed in Fig 3.18a for $H \sim 75 - 100$. Considering that a general increase of the ssDNAs length hinders the enhance of the whole system curvature, see sec. 3.5.2 and Fig 3.12b, the origin of the smaller value of $H_{max}$ observed for the $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ case is revealed. When $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$, $\overline{L}$ does not change until the second hydration range, allowing the curvature of the SAM to increase during the whole hydR-I range, i.e. for $H \sim 0 - 100$. However, when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$, $\overline{L}$ starts to increase within the hydR-I range ($H \sim 75$), as a consequence of a change in the ssDNAs structural tendency from broadening along the SAM to narrowing. This increase of the ssDNAs length at $H \sim 75$ hinders the enhancement of the SAM curvature observed at lower hydration levels ($H = 0 - 75$, see Fig 3.15g), thus explaining why the value of $H_{max}$ is smaller for the new grafting density case ($H_{max} \sim 75$) than in the previous one ($H_{max} \sim 100$).

Finally, to understand the second main difference observed in the evolution of the SAM curvature when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$, i.e. its pronounced reduction between 75 and 100 water/ssDNA, we must consider again that the structural homogeneity grade of the SAM substantially varies during the hydration process. We have already discussed this variation via the evolution of the spatial dispersion of the $D$ along the whole SAM, see Fig. 3.15f. As shown in that figure, the value of $D$ ranges between 10-20Å from 0 to 75 water/ssDNA, indicating that the ssDNAs distribution is extremely heterogeneous in that hydration regime. However, when $H \sim 100$ that dispersion reduces more than a $50\%$, see Fig. 3.15f, which reveals that the SAM becomes more regular at that hydration point. That sharp reduction of the structural heterogeneity grade of the SAM between 75 and 100 water/ssDNA implicates that the structural tendency of all the ssDNAs rapidly changes from being quite diverse to becoming extremely

similar. Therefore, the movement of all the ssDNAs when $H = 100$ is more collective than in the previous hydration values, prompting a more pronounced general response of the SAM at that hydration level. This intensification of the SAM mechanical response due to its sharp homogenization was also observed when $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$, see secs. 3.3 and 3.5, but in this case the ssDNAs structural evolution tends to increase the SAM deflection. Therefore, the explanation for the pronounced $\chi$ reduction observed for the SAM with $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ immediately after the curvature maximum is not only that the heterogeneity grade of the SAM structure sharply reduces, but also that it happens when the deflection of the SAM starts to decrease. With that result, we complete the explanation of the two main differences observed when the grafting density of the SAM is reduced from $4.0\ x\ 10^{13}\ cm^{-2}$ to $3.0\ x\ 10^{13}\ cm^{-2}$. Note that we have been able to properly explain these differences only by analyzing the structural changes that the ssDNAs suffer during the hydration process. This highlights again the importance of that ssDNAs structural changes to fully understand how the SAM mechanical response can be tuned with the hydration and the effect of the SAM grafting density value in that hydration-driven tunability.

## 3.7 Conclusions

In this work, we have studied how the hydration affects to the curvature (mechanical response) of a biosystem consistent on a SAM restrained to a graphene substrate. To that end, we have performed long-MD *all-atom* simulations to obtain thermally equilibrated configurations of this system at different hydration levels, i.e. from 0 to 600 water/ssDNA. We have performed these simulations using two different grafting densities: $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ and $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$. With that simulations, we conclude that the curvature of a SAM can be sharply tuned with the hydration. Moreover, we observe that the hydration driven response of this system depends on its grafting density in agreement with previous experimental results [130, 160]. A detailed structural and energetic analysis of these simulations allows us to put in contact the atomic water-ssDNA interaction with the changes observed in the curvature of

the whole nano-system. The results here obtained highlight the importance of the SAM structural changes arisen from the hydration to completely understand its mechanical response with the hydration and the effect of the grafting density. More concretely, in the $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ case, the high water affinity of the ssDNA strands at low hydration leves, i.e from 0 to 100 water/ssDNA, induces an increase of the number of h-bonds bridges formed between one water molecule and various ssDNA atoms. Thus, the ssDNA strands become stiffer, which causes an enhance of the SAM curvature once the structure of the system is sufficiently homogeneous. When the water-ssDNA affinity starts to decrease, i.e from 100 to 300 water/ssDNA, the water molecules enter into the SAM nanopores not only to interact with the ssDNA strands, but also to interact with other water molecules. Consequently, increasing the size of the nanopores by enlarging the length of the ssDNA strands results more favourable than forming more *water-ssDNA bridges*. The increase of the ssDNAs length causes a progressive reduction of the SAM curvature that only stops at very high hydration levels, i.e when H $\sim 400$ . From that hydration point, the new water molecules added to the system do not enter more into de nanopores but remain on top of the SAM, without practically altering its structure and mechanical response.

For the other SAM, i.e. when $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$, similar results are observed at high hydration levels, but significant differences are detected from 0 to 100 water/ssDNA, i.e. when the high water-ssDNA affinity prompts the formation of more water-ssDNA bridges during hydration. For this grafting density value, the ssDNAs width is smaller than the nanopore size. Therefore, at low hydration levels, the ssDNA strands have to broaden over the surface to approach to its adjacent ssDNAs and then form more h-bonds bridges between one water molecule and different ssDNA strands. This broadening effect increases with the hydration, thus leading to a progressive reduction of the ssDNAs length (the ssDNAs become shorter), which in turn allows the increase of the whole SAM curvature. At H $\sim 75$, the ssDNAs have to narrow again for allowing the entrance of more water molecules into the nanopores, which causes that they stretch with the hydration and then the SAM curvature becomes smaller. This process ex-

plains why the response of a SAM with $\rho = 3.0\ x\ 10^{13}\ cm^{-2}$ does not increases for H $>$ 90, while it is the case for a SAM with $\rho = 4.0\ x\ 10^{13}\ cm^{-2}$ [130, 160]. We expect that the methodology and the mechanisms analyzed here will be essential to undertand not only the changes observed in the mechanical response of ssDNA monolayers with the hydration [130, 160, 161], but also the effect of the water molecules on the structural and mechanical properties of similar systems, such as polymer brushes.

# 4 Tuning structure and dynamics of Azurin junctions via single amino-acid mutation

In the growing field of biomolecular electronics, blue-copper Azurin stands out as one of the most widely studied protein in single-molecule contacts. Its electron-transport properties have been analyzed via experiments based on self-assembled monolayers [187], as well as single-protein wires using Scanning Tunneling Microscopy (STM) [12, 188]. Interestingly, despite the paramount importance of the structure/dynamics of molecular contacts in their transport properties [189], these factors remain largely unexplored from the theoretical point of view in the context of single Azurin junctions. This information might be crucial to a complete understanding of the experimental results, as important Azurin parameters such as its interaction with the electrodes, or its orientation relative to them, can not be straightforwardly gathered from the experiments [190].

In this chapter, we address this issue performing *all-atom* MD simulations of *Pseudomonas Aeruginosa* Azurin adsorbed to a Au(111) substrate. In particular, we focus on the structure and dynamics of the free/adsorbed protein and how these properties are altered upon various single-point mutations. The results revealed that wild-type Azurin adsorbs on Au(111) along two well defined configurations: one tethered via cysteine groups and the other via the hydrophobic pocket surrounding the $Cu^{2+}$. Surprisingly, our simulations revealed that single amino-acid mutations gave rise to a quenching of protein vibrations ultimately resulting on its overall stiffening. Given the role of amino-acid vibrations and reorientation in the dehydration process at the protein-water-substrate interface [139], we suggest that this might have an effect on the adsorption process

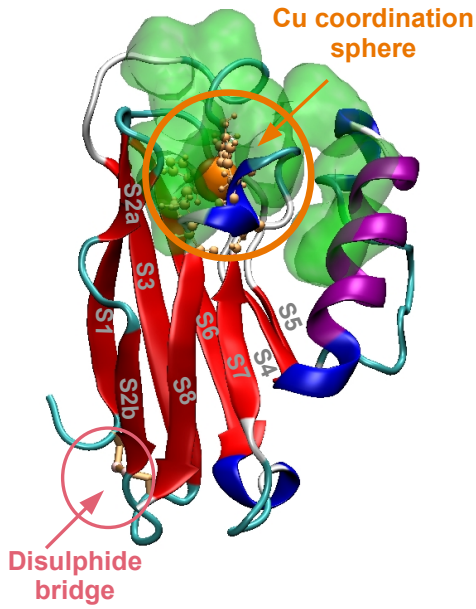



**Figure 4.1:** ***Characteristics of the Azurin structure.*** *The Azurin is represented according to its secondary structure content: eight $\beta$-sheets (red), a single $\alpha$-helix (purple), 310-helix (dark-blue), turns (cyan), and random-coils (white). The copper atom is shown using its vdW representation in an opaque orange color, and its coordination residues are represented with ball-stick model. The disulfide bridge and the main chain of the two cysteines (Cys3,Cys26) which formed it are colored in light orange. The sulfur atoms of these two cysteines are highlight in pink. The position of the hydrophobic patch has been marked with a green Connolly surface [106].*

of the mutant, giving rise to new adsorption configurations.

## 4.1 Introduction

### 4.1.1 Metalloproteins and Azurin

Metalloproteins are characterized by containing a metal ion cofactor bonded to the organic protein peptide chain. That characteristic allows these proteins to maintain unique tertiary structures and to perform electronic transport (redox) processes, which are essential to accomplish the important biological functions where they are involved. These biological functions comprises, for example

the transport of $O_2$ and $CO_2$ through the blood, the oxidative phosphorylation needed for the respiration of animal cells, and the photosynthesis process carried out in plants and certain bacteria [191].

Among the different metalloproteins, Azurins have been a subjet of growing interest in the last years [189, 192–196]. This protein belongs to the cupredoxins family, i.e. its metal ion consists on a copper one. It is a quite small protein, with a molecular mass of 14 kDa and only 128 amino-acids ($\sim$ 2000 atoms). In this protein, the metal ion (the copper) is ligated to five amino-acids of the peptide chain, i.e. the gly45, hys46, cys112,his117, met121. This protein region constitutes the copper coordination sphere and it is assumed to mediate the Azurin electronic transport process, which is characterized by the reversible change of the Cu redox state from Cu(I) to Cu(II). This electronic transport process is related to the denitrification mechanism performed by certain bacteria, such as the Pseudomonas, Bordetella and Alcaligenes [191]. The characterization of the Azurin structure is completed by describing its main secondary structure elements, i.e. a $\beta - barrel$ (formed by eight $\beta - sheets$, see Fig. 4.1) and a solely $\alpha - helix$. The copper coordination region (Cu + 5 ligands) is situated at one edge of the $\beta - barrel$ and surrounded by a set of hydrophobic amino-acids (Azurin hydrophobic patch, see Fig. 4.1). In the other edge of the $\beta - barrel$, the protein posses its unique disulphide bridge (Cys3-Cys26), which can be used to anchor the protein to surfaces as it is quite exposed to the surrounding media.

### 4.1.2 Azurin electron transport experiments: Blinking methodology

In the last decade, the unique interplay between optical and vibronic properties with electron transfer [187, 188, 198–200] together with biorecognition capabilities, makes Azurin proteins very promising candidates for molecular bioelectronic applications. However, the mechanism for electron transport through these systems within solid-state junctions remains, to date, unclear. This has encouraged in the last years the development of different experimental works aiming to understand the Azurin electron transport process. The conductance of this protein has so far been measured by various techniques [12, 187, 201–203], including STM which has been exploited in both the break-junction and

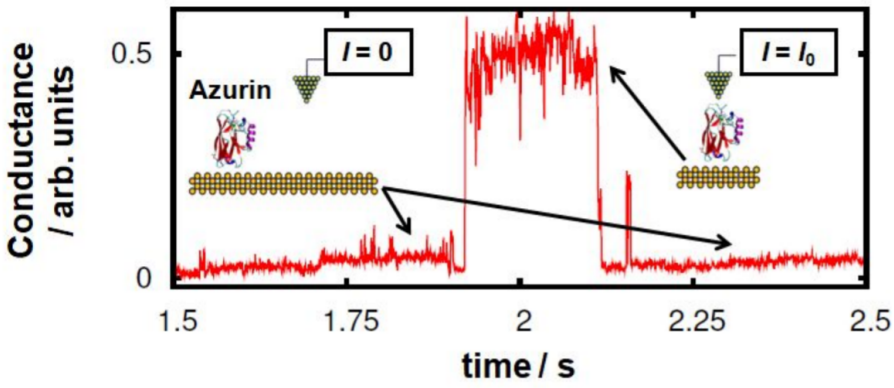



**Figure 4.2:** ***Schematic representation of the blinking experiments.*** *Representative "blink" identified in the experimental transients of the conductance relative to the current flowing between the two electrodes at a constant distance. This kind of event is observed when the Azurin spans the gap between the STM tip and the Au substrate electrodes (see inset). When the protein detaches itself from one of the electrodes, the conductance drops down again to the initial set point level. This figure has been reproduced from ref. 197 by permission of MDPI.*

*blinking* modes [12]. The latter is particularly suitable for measuring proteins, as it allows evaluation of single-protein conductance whilst minimizing disturbance of the protein's secondary structure. In this technique, the tip is brought into close proximity with an individual protein, avoiding mechanical pulling, and kept at a fixed distance from the surface (tunneling distance). The transient current is then monitored. Occasionally, contact is created between the STM tip and the protein due to thermal fluctuations of the molecule, which induce its spontaneous attachment to the tip. This kind of event will be henceforth called "blinking". When this happens, jumps in the current signal are detected, see representative trace in Fig. 4.2. The conductance of the protein can then be measured until the protein detaches from the tip, when the conductance drops again to the initial set point level, see Fig. 4.2. The protein detachment from the tip can happen as a consequence of different factors, such as the tip thermal drift.

This technique was initially applied to small organic molecules [204–207].

Later, the same technique was improved by adding a pulling step subsequent to the blinking event in order to corroborate the presence of a molecule in the junction. This has made it possible to measure the conductance of more complex organic molecules [208] or metal-organic complexes like porphyrins [209]. These measurements have proven highly sensitive to the specific target molecule showing a clear dependence on its size and nature and even to subtle conformational changes or the orientation and the tilting angle of the molecular junction. In addition, it has been demonstrated that this technique can be used to trigger and monitor on-surface chemical reactions [209]. Applying it to small-size biomolecules such as azurins [12], ensures a minimal impact on the general folding protein structure as compared with the STM break-junction technique, where the tip is used for performing a fast pulling of the protein junction. It is important to note, however, that the two techniques have shown quite consistent results on the blue-copper azurin [12], indicating the high mechanical stability of this globular system.

### 4.1.3 Tunability of Azurin electronic properties via single-point mutations

Azurin mutants have also been investigated in depth [210]. The introduction of mutations in the Azurin structure has proven to affect its redox properties [210–212]. More concretely, a highly tunable variability of the protein redox potential, which spans over a 700 mV range, was found, depending on the specific mutant and without significantly altering the redox-active site [211, 212] . This modification of the Azurin redox properties can be achieved not only replacing the nearest residues to the Cu coordination center, but also replacing the outer-sphere residues. Additionally, recent STM *blinking* experiments [12] have evidenced that single amino-acid mutations can affect the electronic transport properties of this metalloprotein. In this experiments, they analyzed the effect of replacing an amino-acid located in the second coordination sphere of the Cu atom (lysine 41) by a cysteine. The introduction of a cysteine has been widely used to increase the stability of proteins between two gold electrodes, as this residue strongly binds to gold (it posses a sulfur atom) [12, 213]. Using the *blink-*

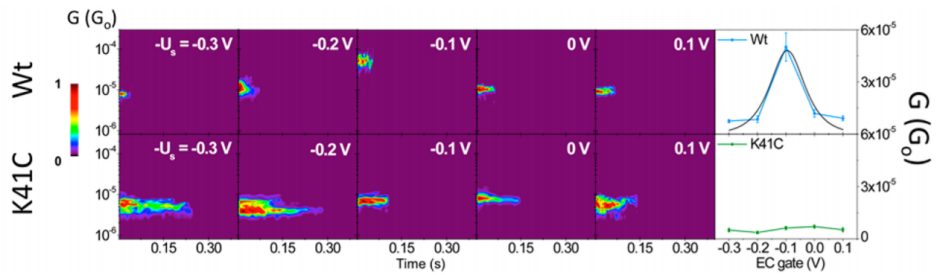



**Figure 4.3:** ***Blinking experiments results: differences in the Azurin electron transport behaviour with the introduction of single amino-acid mutations.*** *2D-blinking maps for both wild-type (top) and K41C Mutant (bottom), proteins at different electrochemical (EC) gate potentials. Several tens (up to a hundred) of individual blinking traces are accumulated to build each 2D map without any selection. The counts have been normalized for each map versus the maximum value so that each 2D map has its maximum count set to 1. The far right graphs summarize the average single-protein conductance (G) vs the EC gate (V) for both studied proteins. This figure has been reproduced from ref. 12 by permission of the American Chemical Society.*

*ing* technique, see sec. 4.1.2, and controlling the EC potential of the system, they were able to measure the conductance of both the wild-type and mutated proteins at different EC gate potentials. The results [12] showed a maximum in the conductance versus the EC gate potential for the wild-type protein, while the conductivity of the mutant did not depend on the EC potential, see Fig. 4.3. This seems to indicate that only changing one aminoacid, the Azurin electron transport process can be drastically changed, demonstrating the feasibility of tailoring charge transport in bioelectrical contacts and opening new horizons in the potential uses of this metalloprotein in biomolecular electronic devices [12].

All these results motivate the understanding of the electronic transport properties of the Azurin on metal junctions and how they can be changed with the introduction of mutations. However, a proper description of Azurin transport mechanisms is extremely complicated, not only because of the complex nature of the proteins but also because of a number of various factors which, to date, remains unclear. Among these factors are the protein interaction with the electrodes [187], the orientation of the protein relative to them [189], the oxidation

state of the Cu coordination center [188, 214] and the role of other moieties different from the Cu center in the transport behavior of this protein [215]. This information is usually difficult to be gathered only from the experiments [190]. Moreover, only few theoretical works [189, 196, 215] have investigated to date the effect of all these factors in the Azurin electronic properties, and only some of them have also analyzed the effect of single-point mutations [215]. One of these works has recently obtained the whole electronic structure of the wild-type Azurin protein and three Azurin mutants with ab-initio calculations, showing that mutations do not introduce any significant change in the Azurin electronic structure [215]. This result suggests that the introduction of mutations may affect to other electronic properties of the Azurin junctions, such as the coupling of the Azurin with the gold electrodes. Considering that this characteristic is very sensible to the protein orientation [189], which in turn depends on the adsorption dynamics of the protein, a better understanding of the structure and dynamics of the Azurin on a metal junction and the role of mutations in these two properties is clearly needed to gain further insight into the tunability of the Azurin electronic transport with mutations [12, 215].

### 4.1.4 Overview

MD simulations of the protein dynamics may allow us to gain further insight into the structure and behavior of the metal-protein-metal junctions [68, 216–218]. These studies can provide information regarding protein orientation, and whether the protein folding is preserved or affected by the interaction with the electrodes [218]. This method has already been used in the past to study the structure and dynamics of the Azurin protein free solved in water [195, 219] as well as anchoring to a gold surface via its cysteines [189, 196, 220]. However, to our knowledge, the systematic analysis of the free adsorption process of this protein onto gold surfaces, i.e. with no preferential anchorage point assumed a priori, has remained elusive. Such free-adsorption study may allow us to unravel not only the characteristics of the Azurin-gold electrode interaction beyond cysteine-gold tethering (e.g. adsorption configurations, protein dynamics, etc.), but also other factors which may influence the protein-surface interaction.

Additionally, to our knowledge, the effect of single-point mutations on the Azurin adsorption process has not been previously studied. Considering that the introduction of these mutations affects to the vibrations/flexibility of the Azurin [25], as well as to the flexibility of other proteins [3, 15, 25–30] (see sec. 1.1.2), their role in the Azurin adsorption dynamics should be inspected to understand their influence on the Azurin electronic transport behavior. For all these reasons, in this work we have decided to perform long MD-simulations (reaching up to 0.5$\mu$s for some of them) to analyze the structure and dynamics of a wild-type Azurin from the *pseudomonas aeruginosa* in solution as well as its adsorption to Au(111). Furthermore, we wanted to explore how the structure and dynamics are affected via the introduction of single amino-acid mutations. We have analyzed the effect of three different mutations based on substituting a single residue of the protein chain (K41, L120, and S89) by a cysteine, i.e. the K41C, L120C and S89C mutations. In this chapter, the results of that study are described.The chapter is organized as follows. In sec. 4.2, the details of the MD simulations used for studying the structure and dynamics of the wild-type and mutated Azurin variants both unrestrained in water and during the adsorption process are explained. In sec. 4.3, the structural and dynamical results obtained during the unrestrained MD simulations in water for both the wild-type and mutated proteins are presented. A detailed analysis of the fluctuations per residue in the three mutated structures considered in this work reveals that the introduction of mutations quenches the flexibility of some turn regions of the protein, leading to an overall stiffening of the Azurin structure. In sec. 4.4, the structural and dynamical results obtained during the MD adsorption simulations for the wild-type protein and one mutated protein (K41C) are detailed. This work suggests that single-point mutations stiffer the Azurin structure, which may influence the dynamics of this protein during adsorption (smaller capability for self-reorienting), thus resulting in different final adsorption configurations than the ones obtained for the wild-type protein.

## 4.2 Methods

### 4.2.1 General MD details

In this subsection, we report the general simulation details of all the MD simulations performed in this work. We have considered five different Azurin proteins: wild-type Azurin, Apo Azurin (without the copper ion), and three wild-type mutants. The X-ray crystallographic structure of wild-type Azurin was obtained from the protein data bank [221] with the PDB code 4AZU [193]. Protons were added to the protein structure according to the calculated ionization states [105] of its titratable groups at a pH of 4.5, in accordance with recent experiments [12]. The Apo initial structure consisted in simply removing the copper ion from the crystallographic structure of wild-type Azurin. The three Azurin mutants here considered have been prepared by replacing a given amino-acid (lysine 41, leucine 120 and serine 89) by a cysteine. This particular mutation is expected to promote the anchoring of the newly added cysteine to the gold contacts [12, 222]. The residue replacement was performed changing the amino-acid type and removing the side-chain of the mutated amino-acid (lysine 41, leucine 120, serine 89) in the wild-type protein PDB with a text editor. The position of the atoms of the new side-chain was selected in agreement with the CYS ligand structure extracted from the protein data bank [221] (see the side-chain conformation of the mutated residues in Fig. 4.4). Note that although all mutations are in the vicinity of the copper(II) ion they are located at different distances from it, see Fig 4.4. In the L120C and K41C, the mutation is located in the second coordination sphere of the Cu atom ($d_{Mut-Cu} \sim 9$ Å) whilst in the S89C, the mutation is in a flexible coil near the Azurin $\beta$-barrel at a distance of $d_{Mut-Cu} \sim 11$ Å from the copper(II) ion (see Fig 4.4). The net charge of the resulting structures is zero for the wild-type, L120C and S89C, and -1 for the K41C and Apo proteins. In that last two cases a $Na^+$ counter-ion was added to neutralize the net charge of the system.

For studying the Azurin adsorption process, we prepared a surface consisted on a Au(111) three atomic layers-thick slab. The initial cell used for creating this surface was a hexagonal cell with the lattice parameter of the Au(111) (2.9Å).

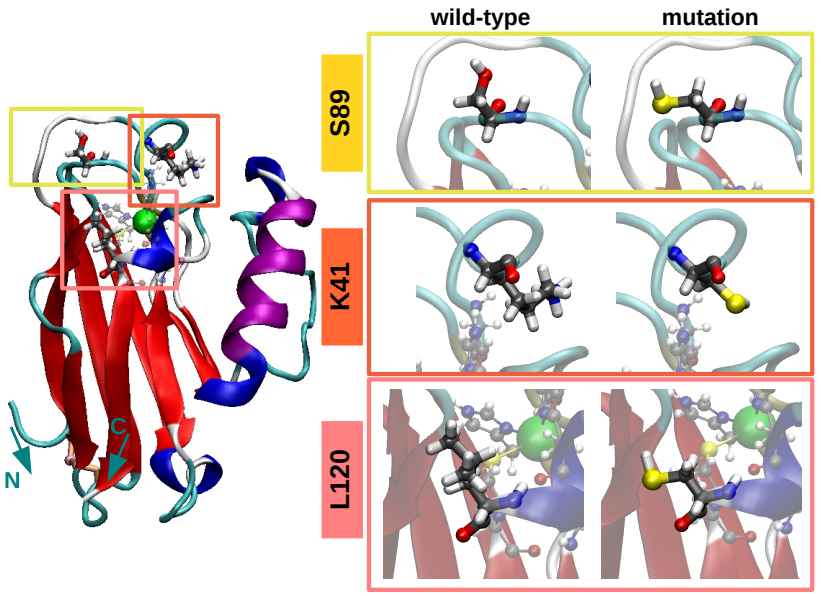



**Figure 4.4:** ***Initial configuration of the Azurin proteins.*** *The Azurin is represented with its secondary structure: $\beta$-sheet (red), $\alpha$-helix (purple), 310-helix (dark-blue), turns (cyan), and random-coils (white). The copper(II) ion is shown using its vdW representation in an opaque green color, and its coordination residues are represented with a ball-stick model. The disulfide bridge and the main chain of the two cysteines which formed it are colored in light orange. The sulfur atoms of these two cysteines are highlight in pink. The position of the three mutated residues (lysine 41, leucine 120 and serine 89) and the initial configuration (prior to minimization stages) of their side-chains in the wild-type and mutated proteins are here also indicated (orange,pink and yellow respectively). The distance between the Cu atom and the alpha carbon of the K41/L120/S89 mutated amino-acid is $d_{Mut-Cu} \sim 9.1$Å$/8.9$Å$/11$Å.*

Once the three-layers slab was created, it was truncated to get a rectangular slab with dimensions $8 \times 8\ \mathrm{nm}^2$ in the x-y directions. A minimization of the surface structure alone was performed before the adsorption simulations to prevent steric clashes. The Au(111) lattice parameter value was not affected during this preliminary simulation. The positions of the atoms in the lowest layer were fixed during the MD runs using a harmonic restraint of 5 kcalmol$^{-1}$.

The ff14SB force field [40] was used to describe all standard amino-acids

present in the Azurin. The inter-atomic potentials of the copper(II) ion and its corresponding 5 ligands were described using a quantum mechanically derived force field [223]. This force-field includes both bonded and non-bonded terms between the Cu ion and its 5 ligands and has been widely used to model the blue-copper Azurin protein [12, 224–227]. In particular recent experiments [225] have shown how early stages of mechanical unfolding of this protein are well described by this force-field. The gold surface was described using the CHARMM-METAL [228, 229] force field, which is thermodynamically consistent with the force field used to describe the protein and has been successfully employed to study similar inorganic-bio-molecular interfaces [229, 230]. Moreover, all simulations were performed in water, which was explicitly modeled using the TIP3P force field [114]. The Joung/Cheatham parameters were used to describe the sodium counter-ions [174, 231].

All the simulations were performed using the AMBER software suite [35] with NVIDIA GPU acceleration [118, 119, 232]. We used PBC with a rectangular box larger enough to prevent the interaction between the protein and its periodic images. PME, with a real–space cutoff of 10 Å , was used to account for long–range electrostatic interactions. vdW contacts were truncated at the real space cutoff of 10 Å . In all the simulations, the temperature of the system was adjusted by means of a Langevin thermostat [48] with a friction coefficient of $\gamma = 1$ ps$^{-1}$ . For the simulations performed in the NPT ensemble (equilibration runs, see simulation protocols subsections below), a Berendsen barostat [50] with a relaxation time of $t_p = 1$ ps was used. The validity of this barostat for equilibrating the system density has been previously confirmed in sec.2.2.2. The SHAKE [122] algorithm was used to constrain bonds containing hydrogen, thus allowing us to use an integration step of 2 fs. Coordinates were saved every 1000 steps

### 4.2.2 Simulation Protocol for the unrestrained Azurin dynamic in water

Unrestrained MD simulations in water were performed for the five Azurin proteins here considered, i.e. wild-type, Apo, K41C, S89C and L120C. The simulation

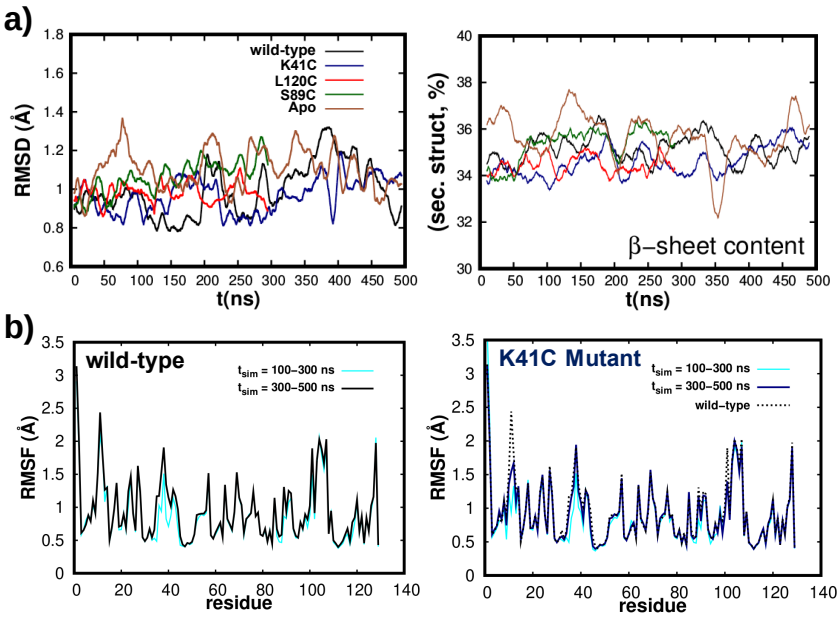



**Figure 4.5:** ***Time-dependency of the Azurin structure and dynamics during its unrestrained simulation in water. a)*** *Time evolution of the RMSD and the $\beta$-sheet content during the unrestrained simulations in water for all the proteins here considered (see more details about the methodology used for computing these two structural magnitudes in sec. 4.2.4). From this figure, we conclude that no major structural changes arise during the last 200 ns of the larger simulations, i.e. the ones corresponding to the Apo, wild-type and K41C mutant proteins. Therefore, 300 ns of simulations are enough for characterizing the structure of these Azurin proteins in water.* ***b)*** *root-mean-square fluctuation (RMSF) per residue of the wild-type (left panel) and K41C (right panel) proteins calculated in two time ranges of their unrestrained simulations in water: $t_{sim}$=100-300 (cyan) and $t_{sim}$=300-500 ns (colored) (see sec. 4.2.4 for details about the methodology used for computing the RMSF). As observed in this figure, the RMSF distribution along the protein is practically identical in these two time ranges for both Azurin variants, thus indicating that 300 ns of simulation are enough for a proper description of the unrestrained Azurin fluctuations/dynamics in water.*

protocol consisted of four stages. First (stage 1), we prepared the system by embedding the protein in water in such a way that the minimum solute−water distance is 1 Å, thus resulting in a system with dimensions $\sim$ 72Å$\times$66Å$\times$70Å. Second (stage 2), we energy minimized the structures using a combination of

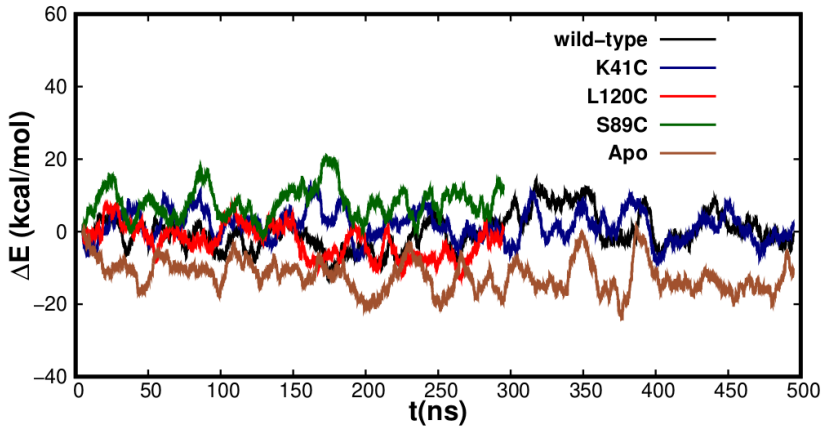



**Figure 4.6:** ***Evolution of the energy of the system during the unrestrained Azurin MD simulations in water.*** *We show the energy evolution obtained for the wild-type (black), K41C mutant (blue), L120C mutant (red), S89C mutant (green) and Apo (brown) proteins. The average of the total energy in the first 10 ns has been taken as our energy reference. The L120C and S89C results go only until 300 ns as these simulations are shorter than the other three.*

steepest descent and conjugate gradient methods to avoid steric clashes. Third (stage 3), we heated up the system from 0 to 300 K with a 2 ns NPT simulation to ensure a well-characterized water distribution at $T = 300$K and $P = 1$ atm. Finally (stage 4), we performed the MD production runs in the NVT ensemble, considering that the system pressure was already stabilized in the previous stage and that the computational cost associated to NVT simulations is lower with respect to NPT. We have simulated the dynamics of the wild-type, Apo and K41C during 500 ns, and the dynamics of the L120C and S89C during 300 ns. We did not continue these last two simulations up to 500 ns (L120C,S89C) as we did not observe any major difference in the evolution of the structure (see Fig. 4.5a) and the fluctuations (see Fig. 4.5b) of the wild-type and K41C proteins from 300 to 500 ns of MD simulation. As shown in Fig. 4.6, the evolution of the total energy of the system was stable during the last $\sim$ 200 ns for the five cases indicating that the simulations are sufficiently long to obtain thermally equilibrated structures.

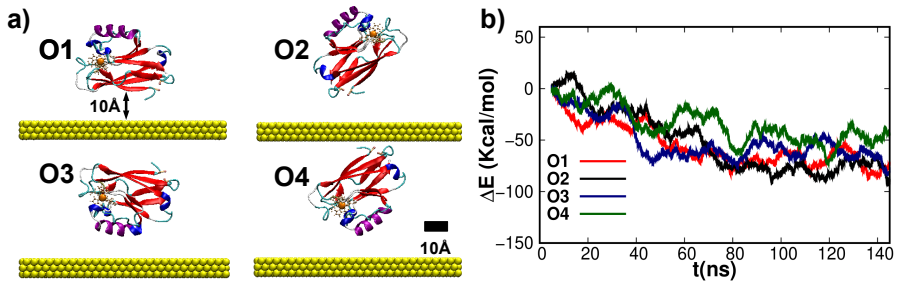



**Figure 4.7:** ***Characteristics of the MD simulations of the Azurin adsorption process.*** **a)** *Representation of the four initial orientations of the wild-type over the whole gold surface considered in this work: O1, O2, O3 and O4. The Azurin is represented with its secondary structure as in Fig. 4.4. The copper atom is shown using its vdW representation in an opaque orange color, and its coordination residues are also represented in orange with ball-stick model. The main chains of the two cysteines are colored in light orange and their sulfur atoms are highlight in pink. The gold atoms are represented with a vdW representation in an opaque yellow color.* **b)** *Time evolution of the total energy variation during the simulation of the wild-type adsorption process for each Azurin-gold initial orientation: O1 (red), O2 (black), O3 (blue) and O4 (green). The average of the total energy in the first 10 ns has been taken as our energy reference.*

### 4.2.3 Simulation Protocol for Azurin Adsorption

We simulated the adsorption process of two different proteins (wild-type and K41C) on the Au(111) slab. The simulation protocol was composed by the same four stages used in the unrestrained simulations mentioned in sec. 4.2.2. The only two differences with respect to the previous protocol are detailed below. *Firstly*, the system preparation (step 1) was not the same. In the adsorption simulations, the Azurin is positioned above the Au(111) surface along four different initial orientations (see Fig. 4.7a). The choice of these orientations is based on the tertiary structure of the protein. Namely, considering the Azurin as a cylinder whose main axis passes through the center of the $\beta$-sheet barrel, then this axis can be oriented standing vertically over the surface (O2, O4, see Fig. 4.7a) or parallel to the surface (O1, O3, see Fig. 4.7a). Moreover, considering the strong interaction between cysteine groups and Au[213, 233, 234], we distinguished between the vertically standing Azurin with the cysteines close

to the surface (O2, see Fig. 4.7a) or further away from it (O4, see Fig. 4.7a). The same distinction holds for the Azurins aligned parallel to the surface, i.e. O1/O3 (close/far) shown in Fig. 4.7a. Regardless of the initial orientation, the initial protein-surface distance is $\sim$ 1 nm, see Fig. 4.7a. This provides enough freedom for the protein to reorient itself if it must prior to its adsorption. The whole protein-surface complex was embedded in water, resulting in a system with dimensions $\sim$ $108\text{Å}\times108\text{Å}\times80\text{Å}$. *The second difference* concerns the duration of the production runs (step 4), as we have simulated the adsorption dynamics for both wild-type and K41C proteins during 150 ns of NVT simulation. As shown in Fig. 4.7b, the evolution of the total energy of the system is stable during the last $\sim$ 70 ns of simulation regardless the initial Azurin-gold configuration used, thus indicating that the simulations are sufficiently long to obtain thermally equilibrated adsorption configurations.

### 4.2.4 Magnitudes analyzed

In order to characterize the protein structural stability we have evaluated the following quantities: the RMSD of the protein backbone [124], the protein $R_g$ [124], and the protein secondary structure content [125, 126]. The RMSD was calculated considering all the backbone atoms of the protein and using the Azurin crystallographic structure as the reference structure. Moreover, to characterize the protein adsorption configurations we have computed the evolution of the CSA between the protein and the gold substrate. The latter is calculated using the following definition:

$$CSA(t) = \frac{1}{2}\left(SASA_P(t) + SASA_S(t) - SASA_{P-S}(t)\right) \ , \qquad (4.1)$$

where the time dependent solvent-accessible-surface-area $SASA(t)$ [106] was calculated for the Azurin protein ($SASA_P(t)$), the Au substrate ($SASA_S(t)$), and the protein-substrate combined system ($SASA_{P-S}(t)$). Additionally, to characterize the fluctuation of the protein during our MD simulations, we have computed in all the cases the RMSF of its residues [124]. That magnitude was computed considering only the protein dynamics that is in energetic equilibrium, i.e. the last 200 ns of the unrestrained Azurin dynamics in solution (see

Fig. 4.6) and the last 70 ns of the Azurin adsorption simulation (see Fig. 4.7). The reference structure used for calculating the RMSF was the energetically-equilibrated averaged Azurin configuration obtained in each simulation. The RMSF was calculated following a two-step procedure: 1)We aligned the Azurin trajectory in energetic equilibrium to its corresponding averaged configuration to evade translational and rotational effects in its RMSF value; 2) We compute the RMSF by comparing the aligned Azurin trajectory with its corresponding averaged configuration.

## 4.3 Unrestrained Azurin MD Simulations in Water: Results

### 4.3.1 Structural characterization of the wild-type molecule and its mutants

In Fig 4.8a, we represent the time averaged configurations of the five Azurin structures (wild-type, Apo, three mutants) obtained in an unrestrained MD simulation in water. These averaged configurations have been extracted using only the data of the last 200 ns of simulation, i.e. when the simulations are in energetic equilibrium (see Fig 4.6 in sec. 4.2.2). The wild-type (black) and Apo (brown) proteins are almost identical to the crystallographic one [193], in agreement with previous results [235]. A quantitative estimation of this similarity is provided through the time evolution of their RMSD shown in Fig.4.8b. There, we find that when the system is in thermodynamic equilibrium (i.e. in the last 200 ns) the standard deviation of the RMSD is smaller than $\sim$ 0.2 Å . Interestingly, this shows that Azurin is a relatively stiff molecule in contrast to the most abundant plasma proteins such as IgG and BSA [56, 97, 236] (see chap. 2), which is certainly a desirable feature for its incorporation in biomolecular solid state devices. Furthermore, we also note that for both structures the RMSD is below 1.4 Å . This shows that both structures are practically identical to the crystallographic one, apart from the natural thermal fluctuations occurring in solution. The structural characterization is completed by computing the time evolution of other fundamental structural properties: the $R_g$, and the $\beta - sheet$ and $\alpha - helix$ content (all shown in Figs. 4.8c-e). Their mean values and the

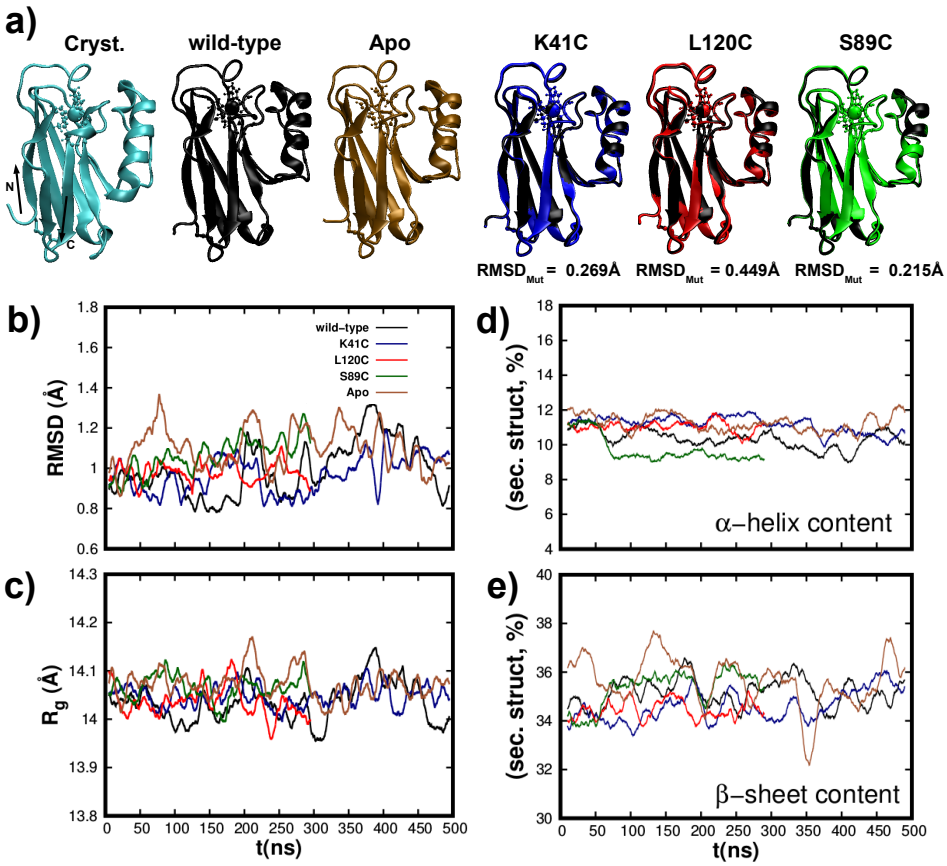



**Figure 4.8:** ***Structural characterization of the five unrestrained Azurin variants in water.*** **a)** *Averaged configuration of the wild-type (black), the Apo (brown), the K41C (blue), the L120C (red) and the S89C (green) Azurin variants are here shown. The crystallographic structure [193] is also shown for comparison (cyan). The Azurin representation used is the same as in Fig. 4.4. Note that the three mutants averaged configurations are aligned with the wild-type averaged configuration (black) and superposed to it with the program Visual Molecular Dynamics. The difference is quantified by the RMSD between both averaged configurations (shown as* $RMSD_{Mut}$*).* **b-e)** *Time evolution of* **b)** *the RMSD,* **c)** *the* $R_g$*, and the percentage of* **d)** $\alpha$*-helix and* **e)** $\beta$*-sheet content for all proteins here considered. Note that the simulations of the L120C and S89C mutants were stopped at 300 ns as this simulation time is enough for getting thermally-equilibrated unrestrained Azurin structures in water (see sec. 4.2.2).*

corresponding standard deviation obtained in the last 200 ns are summarized in Table 4.1. All in all, the obtained values not only support the structural stability of both structures but are also in agreement with previous works [196, 219, 237, 238]. The latter confirms the validity of the force fields here used for describing the structure and dynamic of the wild-type protein with MD simulations.

| Magnitudes | RMSD (Å) | | $R_g$(Å) | | $\alpha$-helix content (%) | | $\beta$-sheet content (%) | |
|---|---|---|---|---|---|---|---|---|
| | mean | $\sigma_t$ | mean | $\sigma_t$ | mean | $\sigma_t$ | mean | $\sigma_t$ |
| **wild-type** | 1.086 | 0.158 | 14.049 | 0.051 | 10.156 | 1.893 | 36.149 | 1.601 |
| **Apo** | 1.099 | 0.137 | 14.061 | 0.062 | 11.151 | 2.128 | 35.299 | 2.002 |
| **K41C** | 1.021 | 0.121 | 14.051 | 0.021 | 10.912 | 1.952 | 35.916 | 1.722 |
| **L120C** | 0.980 | 0.107 | 14.036 | 0.065 | 11.049 | 1.986 | 34.446 | 1.491 |
| **S89C** | 1.100 | 0.110 | 14.062 | 0.060 | 9.319 | 1.532 | 35.674 | 1.463 |

**Table 4.1:** *Mean and standard deviation values of the RMSD, the $R_g$, and the percentage of $\alpha$-helix and $\beta$-sheet content in the last 200 ns of MD simulations for the five Azurin variants (wild-type, Apo, three mutants)*

Concerning the comparison between wild-type protein with its three mutants, i.e. K41C, L120C and S89C, similar results were obtained. As shown in Fig. 4.8a, the mutant structure is practically superimposed to the wild-type average conformation. This suggests that point-like-mutations have little effect on the stability of the protein as a whole, in agreement with previous experimental findings [12]. In fact, this result could be anticipated considering the large amount of h-bonds present in the Azurin molecule, e.g. the ones stabilizing the rigid $\beta$-barrel structure. Computing the RMSD between the wild-type and mutated configurations ($RMSD_{Mut}$, see Fig. 4.8a) we realize that $RMSD_{Mut} < 0.45$Å for the three mutations. This highlights that the difference between the mutant and wild-type proteins is even smaller than the difference between the wild-type configuration and its crystallografic structure arising from thermal fluctuations ($\sim 1.05$ Å , see Fig 4.8b). This result can be understood in light of the position of the mutated residues, which are located at the protein surface (see Fig. 4.4). Consequently, the internal sidechain-sidechain interactions between different amino-acids remain unaltered, which helps to a better preservation of the structure of the protein as a whole [239]. Finally, the structural characterization of the mutants

is completed by computing the time evolution of the RMSD, the $R_g$ and the secondary structure fluctuations (see Fig. 4.8). The average values and their standard deviations are reported in Table 4.1. Interestingly, regardless of the position of the mutation the structural differences between the wild-type protein and its mutants are smaller than the thermal fluctuations (see Table 4.1), indicating that the mutations do not significantly modify the wild-type structure, as observed comparing the averaged configurations. Also the secondary structure analysis shows that the folding of the protein remains virtually unaltered upon the introduction of the mutation (see Table 4.1). This finding is consistent with previous experimental observations, where the structure of other Azurin mutants were analyzed both through their absorption spectra [211] and through their crystallographic structure [212].

### 4.3.2 Dynamics of Unrestrained wild-type and its Mutants in Water

The dynamics of the protein may be accessed via the time averaged fluctuations of each amino-acid composing the Azurin. The averaging was performed in the last 200 ns of MD simulation as then the protein structure/dynamics is in thermal equilibrium. The fluctuations per amino-acid for the wild-type protein (computed as the RMSF, see sec. 4.2.4), are shown in Fig. 4.9a and represented on top of the protein structure in Fig. 4.9b. Additionally, for comparison purposes in Fig. 4.9a we also represent the RMSF values derived from the crystallographic $\beta$-factors [193]. By comparing the results of our simulations with crystallographic data two main features become apparent: firstly the location of the most flexible regions is the same for both MD and crystallographic data, as the peaks appear in the same place in both cases; secondly the fluctuations predicted from the crystallographic data are quenched with respect to our MD simulations. Regarding this last aspect, one must bear in mind that crystallographic data only provides an indicative of the atomic motion/disorder of the protein crystal, whereas the MD simulations also account for concerted oscillations of different parts of the protein due to the presence of the solvent and the temperature of the system [219, 240] . Therefore we may argue that our simulations are able to correctly describe the fluctuations/dynamics of the

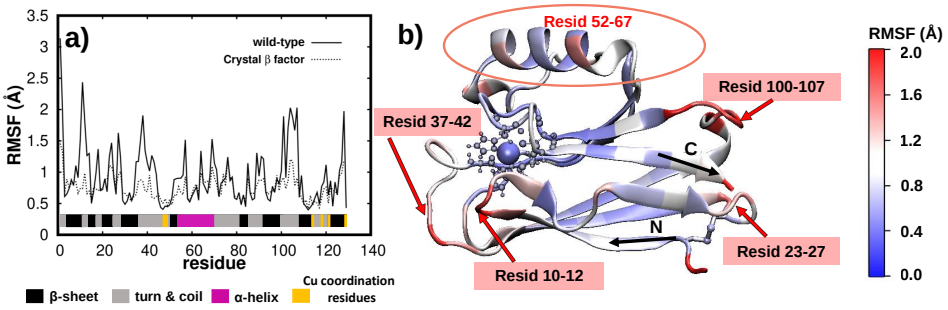



**Figure 4.9:** ***Fluctuations (RMSF) per amino-acid of the wild-type protein during its unrestrained simulation in water.*** **a)** *RMSF of each wild-type residue over the last 200 ns of MD simulation. The same values obtained from the crystallographic $\beta$-factors [193] are also included with a dashed line. Secondary structure of each amino-acid is also detailed (bar below).* **b)** *RMSF data for each residue/amino-acid represented directly over the wild-type averaged structure. The Azurin representation used is the same as in Fig.4.8.*

wild-type protein.

Further insight into the wild-type protein fluctuations, particularly of how they are distributed over the whole molecule, is obtained through their spacial representation as shown in Fig. 4.9b. There, we observe that residues located in *random-coils* or *turns* are generally more flexible (1.0-2.0Å, shown in red in Fig. 4.9b) than the rest. Alternatively, regions with well defined secondary structure held together by strong h-bonds networks such as in $\beta - sheets$ barely fluctuate (0-1.0Å, in blue). These findings are general in nature as similar behaviour is observed in many other proteins [56, 97, 236]. Regarding the Azurin, we observe that its largest fluctuations ( $> 1.8$Å) are located in three different amino-acid segments, all of them belonging to turn regions: residues 10-12, residues 37-42 and residues 100-107 (see Fig. 4.9b). This is also in agreement with previous simulations [25, 219, 220, 241, 242]. It is still worth noting that the only *turn* regions where fluctuations are hampered (i.e. $< 0.8$Å ) are the ones near the five coordination residues of the Cu ion. This shows that the Cu-coordination promotes a rigidification of the neighbouring residues regardless of being located in a soft/hard secondary structure.

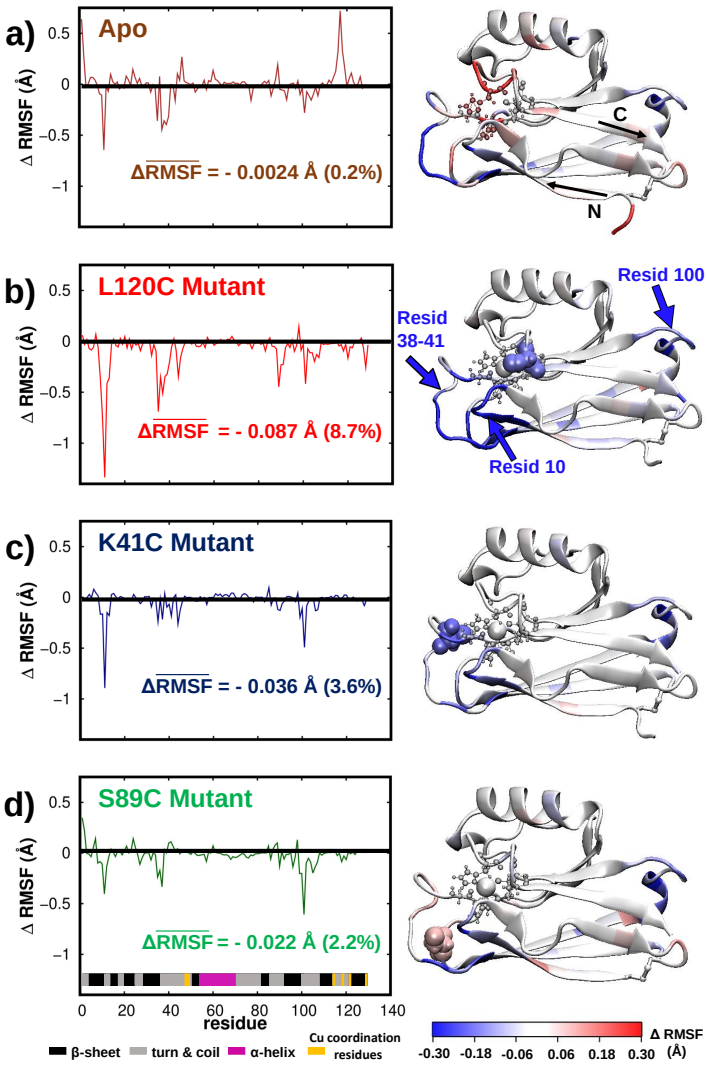



**Figure 4.10:** ***Differential fluctuations of the Apo and mutants with respect to the wild-type protein.*** **a-d)** *In the left panel we plot the* $\Delta\, RMSF(i) = RMSF^{j}(i) - RMSF^{wild-type}(i)$*, being (i) the amino-acid and (j) the different structures considered which are indicated inside the plot as text. Secondary structure of each amino-acid is detailed as in Fig. 4.9a. On the right panel we represent with a color scale these values on-top of the protein's structure. Note that for each protein we also compute the difference between its mean fluctuation value with the mean fluctuation of the wild-type protein, i.e* $\Delta\, \overline{RMSF} = \frac{1}{128} \cdot \sum_{i=1}^{128} RMSF(i) - RMSF^{wild-type}(i)$.

To analyze in more detail the role of the Cu-coordination on the fluctuations/dynamics of the Azurin we now compare the dynamics of the wild-type protein with the Apo. In Fig. 4.10a we represent the difference of the RMSF computed for the Apo with respect to the one computed for the wild-type protein for all amino-acids, i.e. we represent $\Delta RMSF(i) = RMSF^{Apo}(i) - RMSF^{wild-type}(i)$, where i refers to a given amino-acid of the protein. As a result, negative $\Delta\ RMSF$ indicates that fluctuations are smaller in the Apo, and positive $\Delta\ RMSF$ means that Apo fluctuates more. As anticipated, the residues of the Apo located near the copper coordination site vibrate more, with an increase of up to 0.3Å (see Fig. 4.10a). This agrees with previous crystallographic data [243] which showed an increase of the flexibility of the Cu binding site in the Apo structure. Remarkably, the fluctuations in some *turn* regions of the protein (residues 38-40 and 100-103) were also significantly quenched. As a result of this compensation, the mean fluctuation of the whole Apo-protein, defined as $\overline{RMSF} = \sum_{i=1}^{128} RMSF(i)$, is essentially identical to its cooper coordinated counterpart (see Fig. 4.10a).

Similarly to the Apo, we analyzed the role of mutations on the wild-type protein dynamics/fluctuation. Fig. 4.10b-d represents the variation of the RMSF with respect to the wild-type protein for each amino-acid (i.e. $\Delta\ RMSF(i)$) of the three different mutants (K41C, S89C, L120C). Overall, we observe that the average fluctuations are systematically quenched upon the introduction of a mutation (see values of $\Delta\overline{RMSF}$ in Figs. 4.10b-d). Interestingly, the suppression of fluctuations is localized in the three amino-acid segments which have the higher mobility in the wild-type protein, i.e.: residues 10-12, 38-41, and 100-103, all located inside *turn* regions of the protein. From these results three important conclusions may be derived concerning the role of mutations in the dynamics of a wild-type Azurin. First, the introduction of a mutation near the copper(II) ion can affect the dynamics of amino-acids located elsewhere. Interestingly, this long-range effect is in agreement with previous observations [25]. Second, mutations seem to quench the largest fluctuations of the wild-type Azurin regardless the position of the mutation, i.e. their proximity to those regions. Third, the creation of single amino-acid mutations can reduce the mean fluctuation of the

whole wild-type protein up to a $8\%$ , see Figs. 4.10b-d, thus systematically promoting a stiffening of the Azurin. Here it is important to put this observation in a broader context. In many other proteins, e.g. viral capsides, it has been shown that single-point mutations that preserve the protein structure always result in its stiffening [3, 31], see sec. 1.1.2. It has been suggested that this might have some evolutionary origins, as a stiffening of a protein structure may result in a reduction of its biological activity and therefore a natural selection of the more mobile proteins [31]. Thus, the results obtained for the Azurin, i.e. systematic stiffening upon the introduction of mutations, although surprising, are in line with observations on other similar biomolecules [3, 31], see sec. 1.1.2. At last, it is interesting to note that among the three mutations here considered (K41C, S89C, L120C), the largest stiffening is obtained when the mutation is closer to the copper(II) ion, i.e. the L120C structure. In fact, our results indicate that the smaller is the Cu-mutation distance, the more pronounced is the effect of that mutation on the Azurin dynamic, see Figs.4.4 and 4.10b-d. This effect may be understood in-light of the strong coordination of the amino-acids surrounding the Cu site. Whereas in a standard fragment of the protein, a given amino-acid is connect to only two other, around the copper coordination site the amino-acids are connected also with other regions of the protein through the strong metal coordination bonds. Therefore, quenching the fluctuations near this site would have an impact over a larger area of the protein as compared to other not so strongly coordinated region (as is the case of mutant S89C).

## 4.4 MD Simulation of the Azurin adsorption process: Results

### 4.4.1 Structure of the wild-type and K41C proteins upon adsorption tu Au(111)

In the previous subsection, we conclude that the dynamic of the wild-type Azurin protein changes with the introduction of mutations (the protein becomes stiffer). These changes make us wondering if the adsorption dynamics of this protein is also affected by single amino-acid mutations. In this subsection, we answer to this question by analyzing the adsorption process on a Au(111) surface of the

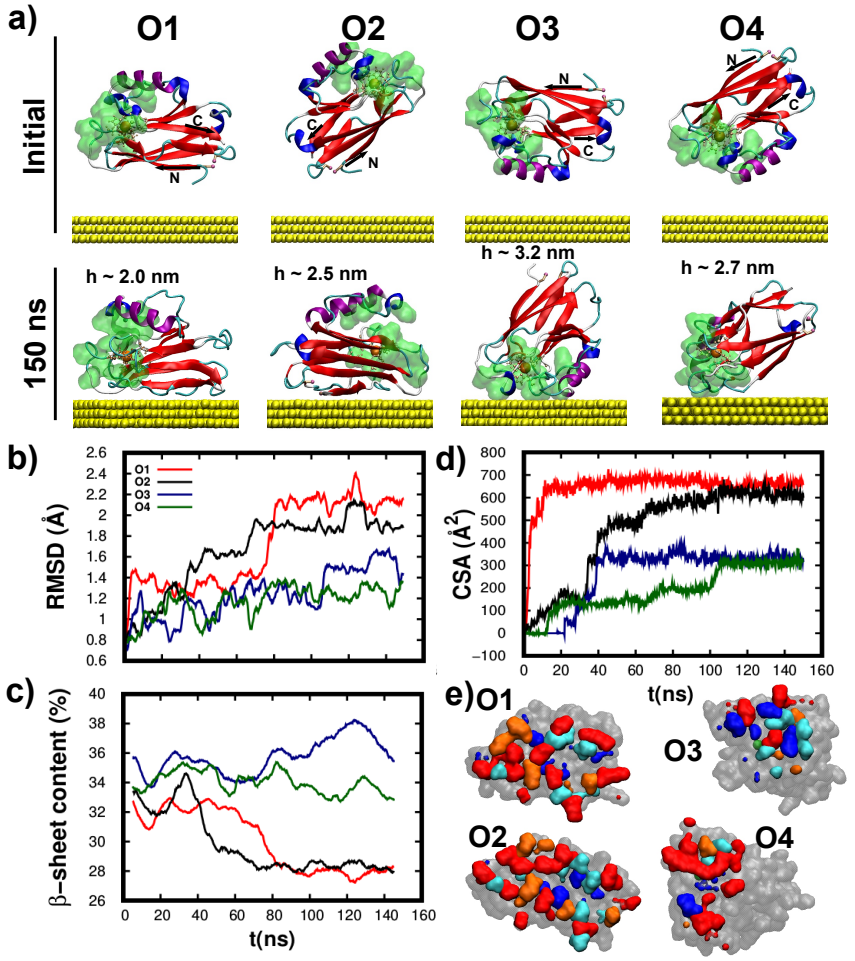



**Figure 4.11:** ***Structural characterization of the wild-type protein adsorption over Au(111).*** **a)** *Snapshots of the initial (top row) and final (bottom row) configurations of the adsorption process. The four different initial adsorption configurations are labeled as O1, O2, O3 and O4. The molecular height (h) of each final configuration is also indicated. The representation used is the analogous to Fig. 4.7. The Azurin hydrophobic patch is highlighted with a green Connolly surface [106]. The Au atoms are represented with yellow vdW spheres.* **b-d)** *Time evolution of the* **b)** *RMSD,* **c)** *β-sheet content and* **d)** *CSA for the four initial orientations considered.* **e)** *Connolly surface [106] of the Azurin amino-acids directly in contact with the Au(111) after the adsorption (t=150 ns). The amino-acids are colored according to their hydrophobicity index: very hydrophobic (blue), hydrophobic (cyan), neutral (orange), and hydrophilic (red). The Connolly surface of the rest of the Azurin residues (non-interacting) is also shown in grey.*

wild-type and K41C proteins with MD simulations. The choice of these proteins is motivated by recent experiments [12] which showed that this particular single point mutation drastically altered Azurin transport behavior (see sec. 4.1.3). The initial protein-surface distance is larger than $\sim$ 1 nm so the protein is able to reorient. Furthermore, four different initial orientations were considered (see the motivation of those in sec. 4.2.3) so to inspect their influence on the final adsorption configuration. Note that to best mimic experimental conditions [12], both the protein and surface are fully embedded in water (see sec. 4.2.3 for further details).

In Fig. 4.11a, we show the final wild-type adsorption configurations obtained for each initial orientation here considered (see sec. 4.2.3). From this figure, we observe that the wild-type protein readily adsorbs over the Au(111) slab (<60 ns) for all starting orientations, in accordance with the fast sample preparation times [12]. Remarkably, as shown in Fig. 4.11a, the protein structure remained almost unaltered upon its adsorption. This is surprising considering that most proteins undergo significant structural rearrangements upon adsorption [56, 97, 139, 236], a process which is more severe for smaller proteins [94] such as the Azurin. This highlights a key feature of wild-type Azurin for its incorporation in solid-state devices, i.e. its strong structural resilience which ultimately may allow to better preserve its activity even in harsh environments. The structural stability may be quantified at the tertiary level via the RMSD (shown in Fig. 4.11b) and at the secondary level through the $\beta$-sheet content (shown in Fig. 4.11c). Comparing these results with the data for protein in solution (see Fig. 4.8b-e), we quantify the small change induced by the adsorption in the protein structure ($\Delta$ RMSD $<$ 1Å; $\Delta$ $\beta$-sheet content <7%). Moreover, we observe that the larger is the change in RMSD and in the $\beta$-sheet content, the larger are the contact-surface-areas between the protein and the surface (CSA – shown in Fig. 4.11d). This suggests that the small structural changes promoted by the adsorption in the wild-type protein are located in the amino-acids sitting at the protein-surface interface. We confirm this hypothesis with a detailed inspection of the protein regions in contact with the surface at the end of the adsorption simulations, see Fig. 4.12. We focus on the two adsorption simulations

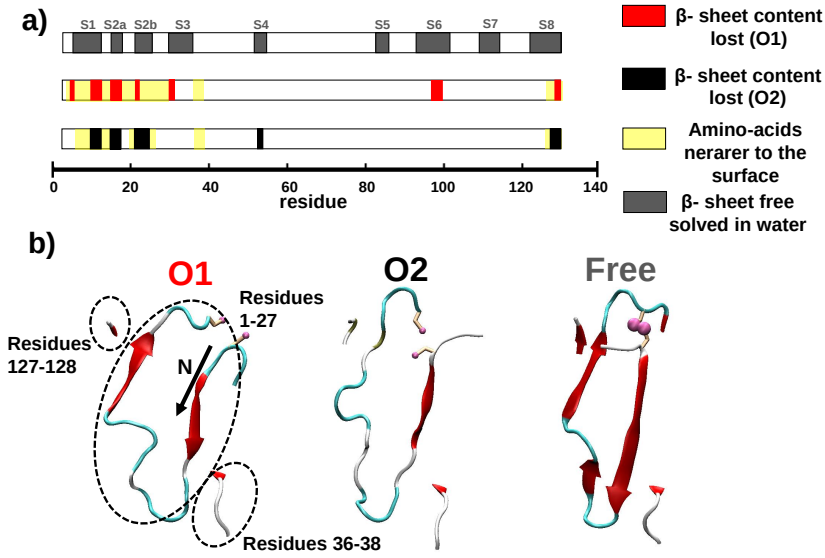



**Figure 4.12:** ***Characterization of the wild-type regions where the*** $\beta - sheet$ ***content is lost during the adsorption simulation.*** **a)** *Amino-acids which have lost its* $\beta - sheet$ *structure at the end of the adsorption simulation for both the O1 (red) and O2 (black) cases. The amino-acids belonging to all the* $\beta - sheets$ *of the Azurin wild-type protein are also here shown for comparison (gray).* **b)** *Secondary structure content of the protein regions that are in contact with the surface (closer than 8Å) at the end of the adsorption process for the O1 (left panel) and O2 (center panel) adsorption configurations. These protein regions correspond to three different amino-acid segments: residues 1-27, residues 36-38 and residues 127-128 (marked with yellow color in* **a)***. The secondary structure content of that amino-acid segments in the free solved state of the wild-type protein is also shown for comparison (right panel).*

where larger protein secondary structure losses are observed, i.e. the O1 and O2 adsorption configurations. From Fig. 4.12, the direct correlation between the structural changes promoted during the adsorption and the number of amino-acids in contact with the surface is confirmed, as most of the $\beta$-sheet content lost during adsorption is located in the protein-surface contact regions for both orientations. At last, we analyze the hydrophobic character of the amino-acids closer to the surface at the end of the adsorption, see Fig. 4.11e. This analysis highlights the importance of the hydrophilic amino-acids in the protein-surface interaction (colored in red), as they represent more than 25% of the total con-

tacting residues regardless the adsorption configuration. This result has been also observed for other proteins [97] and evidences the larger accessibility of these residues to the surface as they are normally solvent-exposed.

The free adsorption simulations shown in Fig. 4.11 points out that wild-type protein does not necessarily adsorb through its cysteine groups, see O3 in Fig. 4.11. To rationalize this finding, other factors known to govern protein adsorption must be considered. On one hand, the strong short-range cysteine–Au interaction is efficiently screened by the many water layers between the protein and surface prior to its adsorption. On the other hand, once the protein-surface contact is reached, the large CSA gives rise to a significant long range non-specific vdW interaction [139], see chap. 2. As shown in our simulations, see Fig. 4.11d, the former is sufficiently large to stabilize an adsorption configuration that is not mediated by the cysteine groups. In fact, a close inspection of the adsorption configurations shown in Fig. 4.11a reveals that wild-type protein preferentially adsorbs along two different orientations: lying-down over the substrate (O1 and O2 – with the two cysteines anchored to Au(111)) and a partially standing up configuration (O3 and O4 – tethered to the surface via the hydrophobic patch). Interestingly, this qualitative assessment is also observed in the quantitative structural analysis as all figures of merit (RMSD, $\beta - sheet$ content and CSA) show two distinct trends. In particular, the lying-down orientation gives rise to a systematically larger CSA (see Figs. 4.11d-e) which in turn results in a larger change of both the tertiary (see Fig. 4.11b) and secondary (see Fig. 4.11c) structures. The two adsorption orientations also result in different molecular heights (h), with $h \sim 2.25$ nm for the lying-down and $h > 2.6$ nm for the partially standing up configuration (see Fig. 4.11a). This is in agreement with previous MD simulations [196] and AFM measurements [244, 245], which also predicted a high variability in the dimensions of the Azurin adsorbed over Au(111). This high variability may be important to understand Azurin junctions conductance experiments, as the protein-electrodes coupling highly depends on the protein orientation [189] . At last, it is interesting to note that for some initial configurations, the protein undergoes a major reorientation process to finally adsorb either through the cysteines or the hydrophobic patch. This is specially

notorious for the O3 case, where the protein rotates over ninety degrees so to finally adsorb via the hydrophobic patch.

The K41C adsorption simulations, see Fig. 4.13a, revealed that it readily adsorbed onto the Au(111) surface. Similarly to wild-type protein, K41C adsorption resulted in minor structural rearrangements of both the tertiary (see Fig. 4.13b) and the secondary (see Fig. 4.13c) structures. Here again, the larger the protein-substrate contact area (i.e. O1 adsorption configuration) the more pronounced were the structural rearrangements (notably the RMSD shown in 4.13b), thus supporting that they are mostly located at the interface. Interestingly, the additional K41C cysteine did not promote cysteine-tethered configurations. This observation supports that the Azurin adsorption process is governed by the detailed balance of many different interactions/processes as aforementioned. Lastly, the as adsorbed K41C molecular heights (h – shown in Fig. 4.13a) also showed a large dispersion consistent with the results obtained for the wild-type protein.

For the wild-type protein, the adsorption proceeded either through the hydrophobic patch or via the cysteine groups, even starting from the O3 orientation when the protein needs a major reorientation to adsorb in one of these two configurations (see Fig. 4.11a). Interestingly, this is not the case for the K41C protein. As shown in Fig. 4.13a, the K41C protein does not significantly reorient itself during adsorption regardless of the initial orientation used, i.e. it adsorbs along its initial orientation. Consequently, each starting orientation gives rise to a different adsorption configuration, see Fig. 4.13a. These may be structurally described as: lying-down with the cysteines in contact with the surface (O1), lying-down with the cysteines far from the surface (O3), upright tethered via the cysteine groups (O2) and upright anchored via the hydrophobic patch (O4). As a result, the final CSA values (see Fig. 4.13d-e) are different for each starting configuration. Interestingly, the two K41C adsorption configurations which were similar to the wild-type case, i.e. lying-down with the cysteines tethered to the surface (O1) and the one anchored via the hydrophobic patch (O4), have a tertiary/secondary structure (RMSD/$\beta - sheet$ cont.) very similar to the one obtained for the wild-type protein.

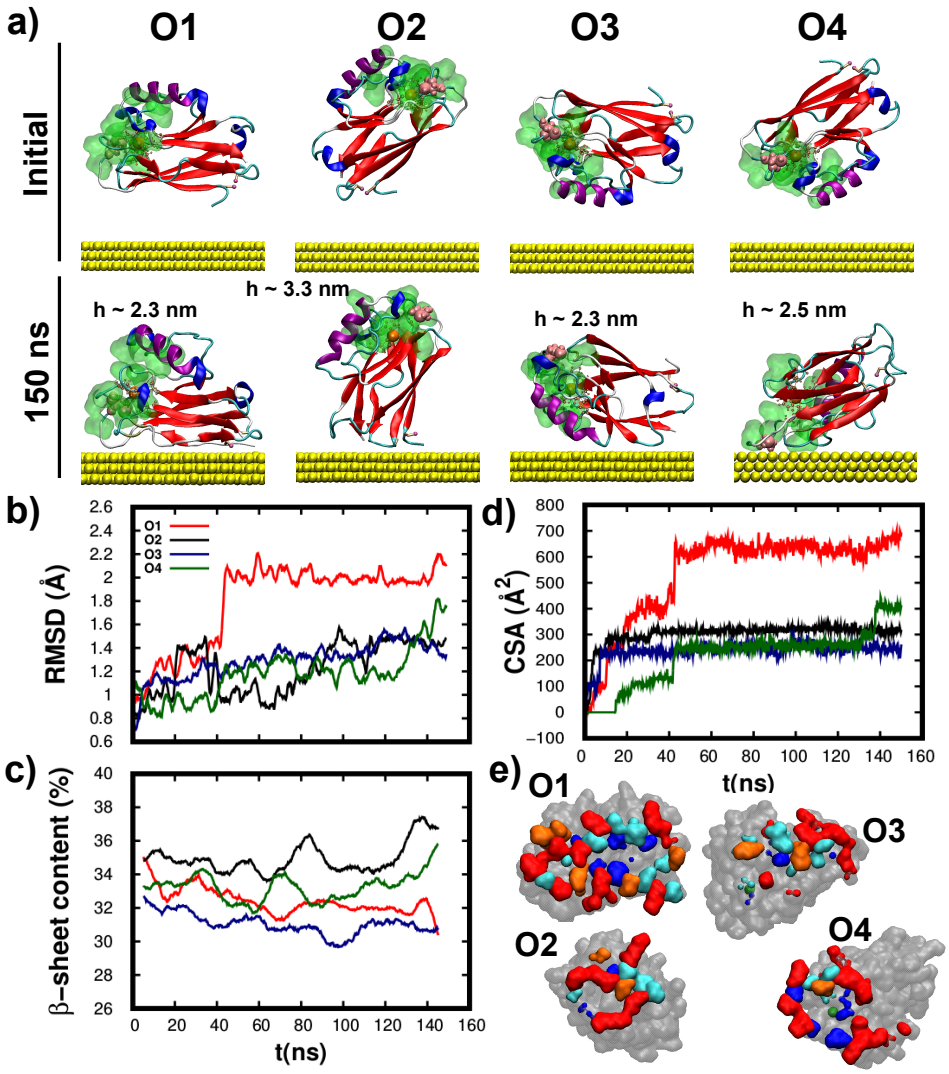



**Figure 4.13:** ***Structural characterization of the K41C adsorption over Au(111).*** **a)** *Snapshots of the initial (top row) and final (bottom row) configurations of the adsorption process. The representation and the initial orientations considered are identical to the wild-type case shown in Fig. 4.11. The mutation site is represented with pink vdW spheres. The molecular height (h) of each final configuration is shown.* **b-d)** *Time evolution of the* **b)** *RMSD,* **c)** *the percentage of β-sheet content and* **d)** *CSA for the four initial orientations here considered.* **e)** *K41C amino-acids directly in contact with Au(111) after the adsorption. The representation is identical to Fig. 4.11e.*

It is worth to mention that despite the undisputed merit of these and other protein adsorption simulations [56, 97, 139], they all rely on the fundamental assumption that a thermal equilibrium has been reached. It is extremely difficult, if not impossible, to determine, from a theoretical study alone, if a MD simulation has reached an actual equilibrium configuration and therefore we do not exclude that other adsorption orientations might coexist as stable/meta-stable states. Possible routes to explore this include: performing longer simulations, additional starting orientations or via the use of accelerated sampling methods. Nevertheless, none would provide a ultimate proof of having reached an actual thermal equilibrium state which can only be provided through high resolution experiments [56], which thus far has remained elusive for Azurin on gold. Concerning the reduction of the protein mobility observed in our simulations with the introduction of a single amino-acid mutation, it is true that our simulations do not provide a direct proof of this effect. Still, they point into the direction that quenching protein dynamics/fluctuations via point-mutations (see Fig. 4.10) might have implications on its adsorption process, even prompting to the variation of the most statistically probable adsorption configurations. This variation of the protein-electrode configuration would influence the electrode-protein coupling [189] , which could ultimately affect to the conductance ability of this protein on the junction [187] . Our hypothesis linking the stiffening of the protein upon mutations and the changes in the adsorption behavior –which we hope that will stimulate future work–, is supported by the results exposed in chap. 2, which show that protein adsorption is ultimately guided by entropic effects at the protein-solvent-surface interface [139] . Quenching fluctuations at a given region of the protein will necessarily hamper its ability to push away the water molecules at the interface via thermal fluctuations and restricts its ability to change the conformation of the amino-acids at the protein-surface interface required to achieve a stable adsorption configuration.

### 4.4.2 Dynamics of wild-type and K41C Mutant upon adsorption to Au(111)

Further insight into the dynamics of the as-adsorbed proteins are provided through the difference between its thermal fluctuations (RMSF) after the adsorption and the RMSF of the wild-type protein in solution, i.e. $\Delta RMSF = RMSF^{adsorbed} - RMSF^{wild-type,free}$. The differential fluctuations per residue over the last 70 ns of simulations are shown in Fig. 4.14 for the four different starting configurations of the two proteins here considered, i.e. wild-type and K41C proteins. Overall, the adsorption process impacts the protein fluctuations/dynamics in three well defined regions: near the surface, far away from it and in the three regions of the Azurin with intrinsic enhanced mobility [25, 219, 220, 241, 242] highlighted in Fig. 4.9. The amino-acids in direct contact with the surface are naturally restrained due to the adsorption process, thus showing much smaller fluctuations that they would have in the liquid phase (i.e. $\Delta RMSF = RMSF^{adsorbed} - RMSF^{wild-type,free} < 0$, blueish regions in Fig.4.14). On the other hand, the residues located far from the surface display an enhanced mobility as shown by their red coloring ($\Delta RMSF > 0$) in Fig. 4.14. The mechanism behind such enhanced mobility results from the fact that the other end of the protein tethered to the surface works as hinge transferring part of the thermal fluctuation energy into global motion most noticed at regions located far from the surface. At last, the amino-acids located in highly flexible regions (residues 10-12, 37-42 and 100-107) experience significant changes in their dynamical properties as a result of the adsorption process. Although small discrepancies among the different orientations, in all cases this effect is felt regardless of the position of these regions with respect to the surface (see first column of Fig. 4.14). Perhaps the most interesting case, since its behavior deviates from the rest, is the O4 orientation for K41C. There, we observe that intrinsically mobile regions, specially the second segment which contains the mutation, experience an enhancement of their fluctuations (corresponding to the positive peaks in Fig. 4.14d), thus deviating from the general trend of quenching fluctuations in these regions as a result of the adsorption. The first aforementioned enhanced mobility peak may be understood in light of a major

rearrangement of the Cu coordination sphere arising from the adsorption of the mutated cysteine onto Au(111).

Surprisingly, Fig. 4.14a-b shows that *equivalent* adsorption configurations of the wild-type protein may result in different dynamical behavior. This is particularly notorious in the $\alpha$-helix region where in the O1/O2 configuration it remained stationary/mobile. Such difference will certainly affect the stability of molecular contacts specially in 'blinking' experiments [12], where contact may be reached further away from the surface and with a higher frequency when the $\alpha$-helix exhibits a mobile character (moving closer and away from the surface) such as in O2 case [197]. As shown in Fig. 4.15a, for O1 the $\beta - sheets$ nearer to the $\alpha - helix$ (S7 and S8 $\beta - sheets$) [219] are more strongly adsorbed to the Au(111) which then restrains the $\alpha$-helix motion (shown in Fig. 4.15c). Contrastingly, in the O2 orientation a slight roll of the $\beta$-barrel equilibrium configuration over the surface gives rise to a larger distance between the surface and the S7 $\beta - sheet$, the S8 $\beta - sheet$ and the $\alpha - helix$, see Figs. 4.15b-c This in turn translates into a smaller interaction of these sites with the surface which ultimately results in a much larger mobility of the $\alpha - helix$ as shown in Fig. 4.14b. All in all, this highlights the interplay between fine structural detail of the as-adsorbed protein and its dynamical properties which will certainly affect the stability of biomolecular contacts.

Concerning the mobility of K41C, we now focus on the orientations which shown an overall decreased mobility as compared to the wild-type protein, i.e. O2 and O3, as they are unable to reorient itself during the adsorption process (see Figs. 4.11 and 4.13). In Fig. 4.14 we observe that in both orientations the regions located near the mutation site (residue 41) experience a strong quenching of their fluctuations, or equivalently are significantly stiffer than the corresponding wild-type protein segments (colored in blue – $\Delta RMSF < 0$). Interestingly, in both cases this region is located far away from the surface, i.e. at a distance larger than 2 nm, pointing out that this quenching of fluctuation is not related to the adsorption process but is instead a direct consequence of the mutation. Since the protein adsorption process is essentially driven by entropic forces and vdW forces (directly dependent on the contact area, see chap. 2) [139]

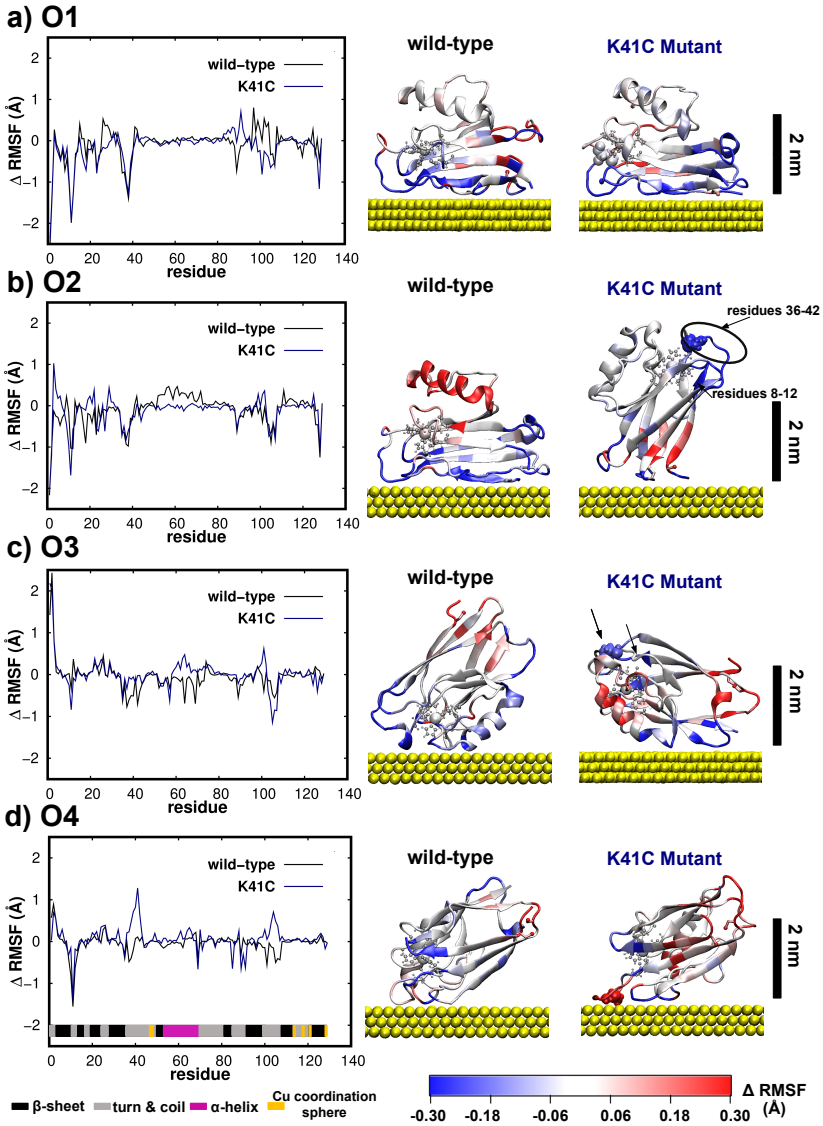



**Figure 4.14:** ***Differential fluctuations of the wild-type and K41C proteins adsorbed with respect to the ones of the wild-type protein in solution.*** *Here we show the* $\Delta RMSF = RMSF^{adsorbed} - RMSF^{wild-type,free}$ *averaged from the last 70 ns of simulations for the four initial orientations* **a)** *O1,* **b)** *O2,* **c)** *O3 and* **d)** *O4. On the left column in black/blue we represent the* $\Delta RMSF$ *value per amino-acid of the wild-type/K41C . The secondary structure character of each amino-acid when the wild-type protein is free solved in water is also specified (see Fig. 4.10). In the right column we show* $\Delta RMSF$ *projected in the equilibrium adsorption configuration. The position of the K41C mutation is highlighted with a vdW representation. The Azurin color representation is the same as in Fig.4.10.*

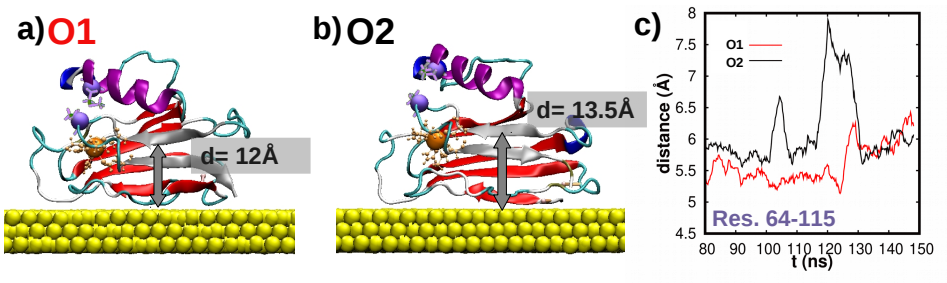



**Figure 4.15:** ***Structural differences between O1 and O2 equilibrium adsorption configurations for wild-type Azurin.*** *Final snapshots obtained for the* **a)** *O1 and* **b)** *O2 orientations. The representation used is the same as in Fig. 4.11. The* $\beta - sheets$ *near the* $\alpha - helix$ *(S7 and S8* $\beta - sheets$*, see Fig. 4.1) [219] are highlighted in gray. The distance between these* $\beta - sheets$ *and the Au(111) surface is also shown. The distance between the Cu center (residue 115) and the end atom of the* $\alpha - helix$ *(residue 64), both highlighted via a vdW sphere in both* **a)** *and* **b)***, is represented in the plot shown in* **c)** *for both orientations, i.e. O1/O2 in red/black.*

it is sensible to assume that soft regions which are able to accommodate larger deformations as a result of the adsorption process will be favored over rigid ones. Therefore, these results seem to suggest that single-point mutations not only affect the overall protein mobility and stiffness but this in turn may also affect equilibrium adsorption configurations. Although our simulations do not provide a direct validation of this effect, we hope that our results might stimulate future works, such as high-resolution AFM images of Azurin on gold, to conclusively prove/disprove this hypothesis.

## 4.5 Conclusions

In this work, we firstly studied how single-point mutations (K41C,L120C,S89C) affect the structure and dynamics of the wild-type protein in solution via all atom MD simulations. The mutations here considered barely changed the wild-type protein structure. Interestingly, they systematically resulted in a stiffening of the structure as a whole. Although, similar findings were already reported in other biomolecules [3, 31] (see sec. 1.1.2), this effect was not previously reported/considered in the Azurin. Such modification of the protein dynami-

cal/mechanical properties certainly plays a relevant role in their stability in solid-state junctions. In the second part of this work, we considered the role of such mutations in the structure and dynamical properties upon their adsorption to a Au(111) surface. Our adsorption simulations showed that the protein structure remained almost unaltered upon its adsorption. Such structural resilience differentiates Azurin from the most common blood plasma proteins [56, 97, 139, 236] where larger changes in both tertiary and secondary structures are expected, specially considering the small size of this protein. Also, our free adsorption simulations revealed a surprisingly high mobility of the wild-type protein which was able to undergo major reorientation maneuver during the adsorption process. Interestingly, such mobility allowed the wild-type protein to systematically adsorb along one of two favored conformations: either with the hydrophobic patch facing the Au(111) surface or tethered by the cysteine groups to the surface. However, for the K41C mutant, our four different adsorptions simulations showed that the Azurin was unable to reorient itself. This resulted in a change of the preferred Azurin configurations leading to the emergence of other two possible adsorption configurations (standing up with the cysteines in contact with the surface, lying-down with the cysteines far from the surface). Given the importance of thermal fluctuations in the adsorption process [139], here we hypothesize that quenching them through the introduction of mutations might ultimately affect how the Azurin adsorbs over Au(111). Although direct validation of this hypothesis remained elusive, we hope that the indirect evidence provided with this work might motivate future works. These results shed light on how two fundamental properties (structure and dynamics) of biomolecular contacts may be tuned via single-point mutations, both known to have major implications in the electron-transport properties of these contacts [189].

# 5 Tuning the graphene thermo–mechanical properties via defect creation

Besides its unique electronic properties [246] , graphene also exhibits a complex thermo-mechanical behavior [247]. One of its most striking properties is its negative TEC [248–250]: contrary to most materials, that expand when heated, graphene contracts. This contraction is due to the thermal excitation of low energy out–of–plane vibration modes. These flexural modes have been reported to govern the electronic transport [222] as well as the elastic response [251] of suspended graphene. In this chapter, a novel experimental technique based on AFM nanoindentations has been used to measure the tension induced by heating in suspended graphene due to the TEC mismatch with the substrate. With this technique, we show that graphene TEC can be reduced by the controlled introduction of low densities of mono-vacancies , reaching a one order of magnitude reduction when the defect density is $5 \times 10^{12}$ cm$^{-2}$. To understand this effect, we also analyze this TEC reduction through atomistic MD simulations. Our MD simulations reproduce the experimental trend and show that TEC reduction is due to the suppression of out–of–plane fluctuations mediated by defects. They also reveal that this suppression is mainly caused by the strain fields created by mono-vacancies in the surrounding areas. These results highlight the key role of defects in the properties of "real-life" graphene and pave the way for future proposals of electronic and mechanical defect engineering. The experiments presented in this chapter were performed by Dr. Guillermo López-Polín, Dr. Cristina Gómez-Navarro and Prof. Julio Gómez Herrero of Universidad Autónoma de Madrid.

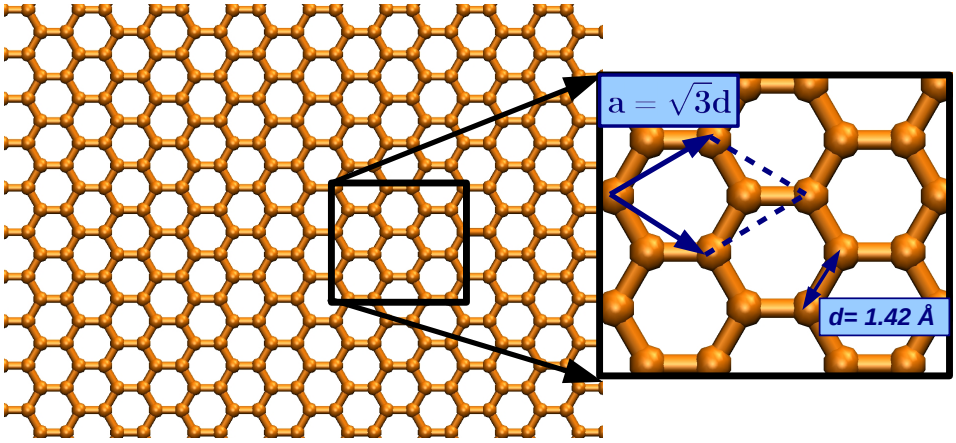



**Figure 5.1:** ***Graphene Structure.*** *Schematic top view of a two-dimensional graphene layer, showing its characteristic honeycomb structure. The primitive cell, the value of the lattice parameter (a) and the nearest neighbor lattice distance (d) are here also shown.*

## 5.1 Introduction

### 5.1.1 Graphene and its thermal expansion coefficient

Graphene is a purely 2D material entirely constituted by carbon atoms covalently bonded and arranged in a honeycomb lattice, see Fig. 5.1. That lattice is described with an hexagonal primitive cell characterized by two carbon atoms and a lattice parameter of $a = \sqrt{3} \cdot d$, where $d = 1.42$ Å refers to the distance between neighboring atoms, i.e. the bond length.

Graphene was isolated for the first time in 2004 [252], becoming the first 2D crystal proven to be stable in its free state. From that date, multiple efforts of the scientific community have been directed to study this peculiar low-dimensional material, observing that it possesses striking properties. Its thermal conductivity is the highest ever measured ($\sim 5 \cdot 10^3\ W \cdot m^{-1} \cdot K^{-1}$) [253]. Moreover, its tensile strength is about $130\ GPa$ and its Young's modulus is near $1\ TPa$, indicating that it is one of the strongest material ever analyzed [254]. These properties have emphasized the promising role of this material in future applications, such as electronic circuits or sensing [255]. For all these applications, the

ability to control its behavior with the temperature is a matter of paramount importance.

One of the most interesting thermal properties of graphene is its thermal expansion coefficient (TEC), as contrary to most materials it is negative, i.e. graphene contracts when heated [46, 248–250, 256–260]. This unusual property arises from the extremely low bending rigidity of graphene ($\kappa_{graphene}$ ∼1.2 eV) [261], which causes that this material exhibits significant out-of-plane thermal fluctuations at ambient conditions [259]. This flexural out-of-plane fluctuations have an strong anharmonic character [259] and their amplitude rapidly increases with temperature [259], thus inducing that the graphene 2D layer shrinks in plane when heated. That in-plane contraction is sufficiently large to counter the usual increase of the inter-atomic distances with raising temperature, thus translating into an actual contraction of the whole material (negative TEC). It is worth mentioning that these out-of-plane thermal fluctuations do not only affect the TEC, but also other graphene properties, as they have also considerable impact on its elastic [4, 247], conformational [262], electronic [263, 264] and thermal transport [265] properties.

Technologically, this graphene thermal contraction is a double–edged sword. It can be exploited in lowering the TEC of regular materials to obtain ultra-stable composites [266] for their use in high precision instrument operating at variable temperature. However, at the same time it is the responsible of device failure when integrating graphene in large scale devices, as the stress induced by the TEC differential at the interface of different materials influences their performance [267, 268]. Therefore, methods to control the TEC of graphene are highly desirable. This interest has encouraged in the last decade the development of both experimental and theoretical methodologies to quantify and tune this graphene property. Experimentally, measuring graphene TEC is a challenging task since conventional experimental techniques are usually designed for bulk materials and cannot be directly applied to such thin membranes. Moreover, the anchoring of one–atom–thick membranes brings out technological difficulties [269]. Therefore, only few experimental works have estimated the pristine graphene TEC value ($\sim \ -7.5 \times 10^{-6}$ K$^{-1}$) [248, 250]. Theoretically, the

graphene TEC has been estimated following different methodologies, such as analytic approaches based on different approximations [256, 259, 270] or atomistic simulations [46, 271], obtaining pristine graphene TEC values ranging from $\sim -3\times10^{-6}$ K$^{-1}$ to $\sim -10\times10^{-6}$ K$^{-1}$. In this work, we aim to estimate the pristine graphene TEC both experimentally and with "all-atom" MD simulations and analyze how it can be tuned with the introduction of defects. It is worth to mention that the TEC absolute value (hereafter referred as |TEC|) obtained from atomistic simulations is usually smaller than the one obtained experimentally [46, 248]. This result can be understood in light of the limited size of the graphene membranes in the simulations. This characteristic introduces an effective cutoff to the long-wavelength out-of-plane graphene fluctuations, which in turn translates into the reduction of the estimated graphene negative TEC [46].

### 5.1.2 Influence of defects in graphene mechanical properties

It has been proven that the unavoidable defects appeared in graphene during synthesis or processing deviates its properties from those of its high perfection form, i.e. pristine graphene [22, 23]. In those works, they found that the graphene Young's modulus of defective graphene membranes was smaller than the one of pristine graphene, see sec. 1.1.2. However, their main objective was not to analyze the defect influence on the graphene mechanical properties and consequently they were not able to directly control the characteristics and number of defects introduced on graphene.

In 2014, Lopez-Polin et al. [4] developed a novel experimental technique to analyze the influence of defects on graphene mechanical properties in a controlled manner. That technique was based on AFM nanoindentations and the irradiation of graphene samples with a known dose of $Ar^+$. With that experiments, they reported an increase in the graphene Young's modulus of $\sim 70\%$ for a dilute density of single–atom vacancies [4], see Fig. 5.2a. In order to address this counterintuitive result, these authors suggested, without further proof, that the induced defects ironed the thermal fluctuations out, see Fig. 5.2b, unveiling the bare Young's modulus of graphene. The base from that suggestion was in the thermodynamical theory of membranes, which predicts that the graphene

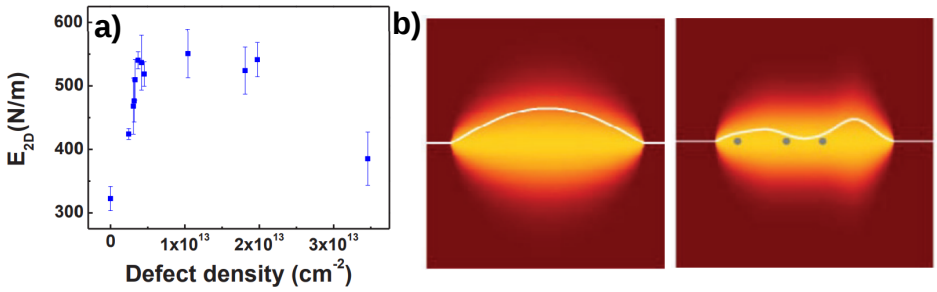



**Figure 5.2:** ***Experimental measurements of the graphene Young modulus at different defect densities performed by Lopez-Polín et al. in ref. [4] a)*** *2D Young's Modulus value ($E_{2D}$) as a function of the defect concentration measured after each irradiation dose in a representative graphene drumhead.* ***b)*** *Schematic representation of the hypothesis suggested in ref. [4] for explaining the effect of defects on the graphene elastic constants (quenching of long-wavelength out-of-plane thermal fluctuations). These figures have been reproduced from ref. 4 by permission of Springer.*

out–of–plane thermal fluctuations (see sec. 5.1.1) should renormalize its elastic constants, making they softer as the amplitude of the fluctuations grows [247, 272] . The introduction of defects supposes a new scale factor for this out-of plane fluctuations, prompting that they become smaller and then the graphene Young's Modulus becomes larger (graphene becomes stiffer). This scenario implies that, not only the Young's modulus, but all the elastic constants and the TEC should be affected by the presence of defects.

### 5.1.3 Overview

Graphene TEC is an ideal candidate to explore the interplay between its intrinsic thermal vibrations and defects. The influence of this interplay on graphene TEC has only been previously investigated theoretically, both by substituting a carbon atom with boron or nitrogen [270] and by introducing mono-vacancies in the graphene membrane [271]. In the first work, they observed that graphene contracts less with the temperature with the introduction of defects, but in the second they concluded the opposite. This controversy reflects that the effect of defects in the graphene out–of–plane fluctuations and more concretely in its TEC is not entirely clear. Considering that this effect is key to understand the physics

underlying many temperature effects in "real-life" graphene, such as the behavior of nanoresonators [273] or the diffusion of water nanodroplets [274], we decided to analyze in detail the TEC of pristine and defective graphene membranes both with experiments and MD simulations. In this chapter, the results of that study are described. The chapter is organized as follows. In sec. 5.2, the details of the experiments and MD simulations used for estimating and understanding the role of defects in graphene TEC are described. In sec. 5.3, the experimental graphene TEC value obtained by our collaborators and its dependency with the mono-vacancy density are detailed. That experiments show that the creation of vacancies on the graphene membranes reduces the |TEC|. In sec. 5.4, we confirm that trend by estimating the TEC of both pristine and defective graphene membranes using NPT-MD simulations. Moreover, we analyze the relation of |TEC| reduction with the modification of the graphene out-of–plane fluctuations. Finally, in sec. 5.5, we try to understand |TEC| reduction by studying with NVT-MD simulations how the strain field in the graphene membrane changes with that mono-vacancy creation. This work reveals that the introduction of a low-density of mono-vacancies reduces the graphene |TEC| due to the quenching of out-of-plane fluctuations caused by the strain fields created by mono-vacancies.

## 5.2 Methods

### 5.2.1 Experimental TEC measurements

In this subsection, we detail the experimental set up used by our collaborators to measure the graphene TEC. The graphene samples were prepared by mechanical exfoliation of natural graphite on Si$O_2$/Si substrates with predefined circular wells of size 1–2 $\mu$m. That circular wells were obtained by patterning with optical lithography and reactive ion etching. AFM images in non-contact mode showed that each graphene flake covers several of these circular wells, forming graphene drumhead structures, see Fig. 5.3b. Thus, these samples allow analyzing the behavior of suspended graphene, as graphene on top of a circular well can be considered as free-standing, see Fig. 5.3a. Only the graphene membranes showing a flat and featureless surface, i.e. absence of bubbles or wrinkles, were

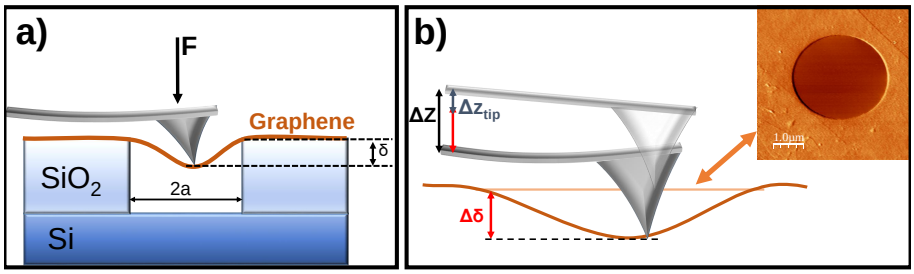



**Figure 5.3:** ***Schematic representation of the experimental set up.*** ***a)****Representation of the AFM indentation experiments over a graphene membrane suspended on a $SiO_2$/Si substrate.* ***b)****Schematic representation of the difference between the indentation ($\delta$) and the cantilever position ($Z$). An AFM image of a representative graphene drumhead is also shown.*

selected for analyzing the TEC.

These suspended graphene membranes were analyzed by indenting a tip at the center of the clamped area with an AFM cantilever, see Fig. 5.3a. That situation can be modeled as a circular membrane with a central point load where the force versus indentation curves behave as [276]:

$$F(\delta) = \pi\sigma_0\delta + \frac{E_{2D}}{a^2}\delta^3 \tag{5.1}$$

where $F$ is the loading force, $\delta$ is the indentation at the central point, $a$ is the drumhead radius, $E_{2D}$ is the 2D elastic Young's modulus of the membrane, and $\sigma_0$ is the stress of the membrane. Experimentally, that $F(\delta)$ curve can not be directly obtained as the cantilever movement is not controlled via its indentation variation ($\Delta\delta$) , but from the variation of its position ($\Delta Z$). That two magnitudes are related by the equation:

$$\Delta Z(F) = \Delta\delta(F) + \Delta\ z_{tip}(F) \tag{5.2}$$

where $\Delta\ z_{tip}$ refers to the deformation suffered by the cantilever during the indentation process, see Fig. 5.3b. To estimate that cantilever deformation at each force value, a calibration measurement is performed consisting on a $F(Z)$ curve over the Si$O_2$/Si substrate, see Fig. 5.4a. In that calibration measurement, $\Delta\delta \sim 0$, which implicates that $\Delta Z(F) \sim \Delta\ z_{tip}(F)$. Then, $\Delta\ z_{tip}(F)$ can be

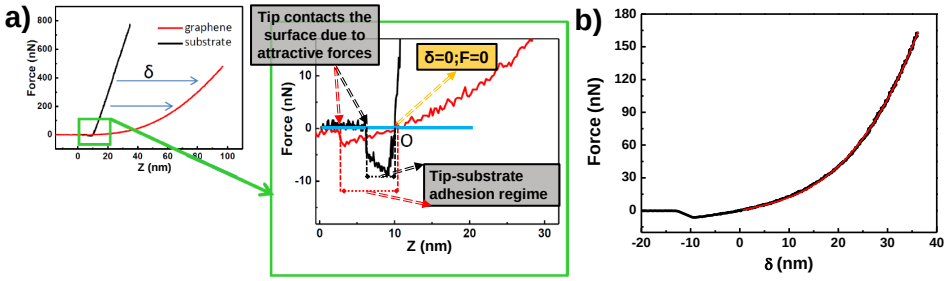



**Figure 5.4:** ***Explanation of the methodology followed for acquiring the experimental*** $F(\delta)$ ***curves.*** ***a)****Representative Force vs.* $Z$ *curves obtained directly from the experiments over the suspended graphene (red) and over the* $SiO_2$*/Si substrate (black). The latter corresponds to the calibration measurement performed to estimate* $\Delta\ z_{tip}$*. We also show a zoom of these two curves in the region of zero force level to highlight the zero displacement point (in yellow). That indentation curves were measured with a Nanotec commercial microscope and* $WSxM$ *software package [275].* ***b)****Force vs. indentation (*$\delta$*) curve measured at the center of the suspended area on a circular graphene drumhead (black). Fitting to eq. 5.1 is overlaid in red (*$\sigma_0 = 0.3$ *N/m,* $2a = 750$ *nm,* $E_{2D} = 385$ *N/m).*

extracted from that calibration measurement and $\delta(F)$ can be computed as the difference between the $\Delta Z(F)$ value obtained on the $F(Z)$ curve performed over graphene and the one obtained in the calibration measurement, see Fig. 5.4a. In that difference, the $Z$ must be identical in both curves when $F = 0$, i.e. when the tip is in contact with the graphene membrane but it is not exerting any load to it, see Fig. 5.4a. Thus, the value of $\delta = 0$ coincides with the point when $F = 0$ and experimental $F(\delta)$ curves can be extracted from AFM indentations. The resulting $F(\delta)$ curves can be fitted to eq. 5.1 by least square minimization, see Fig. 5.4b, with $\sigma_0$ and $E_{2D}$ as fitting parameters. This procedure enables to extract the value of the graphene Young's modulus ($E_{2D}$) and the value of the stress of the graphene membrane, $\sigma_0$. The values obtained in more than 10 drumheads for graphene $E_{2D}$ (350±30 N/m) and $\sigma_0$ (0.2±0.1 N/m) are in agreement with previous works [254, 277].

Towards thermo–mechanical characterization of graphene suspended layers, sample temperature (T) was varied between 283–348 K and consecutive $F(\delta)$ curves were acquired at intermediate temperatures. Upon annealing, graphene

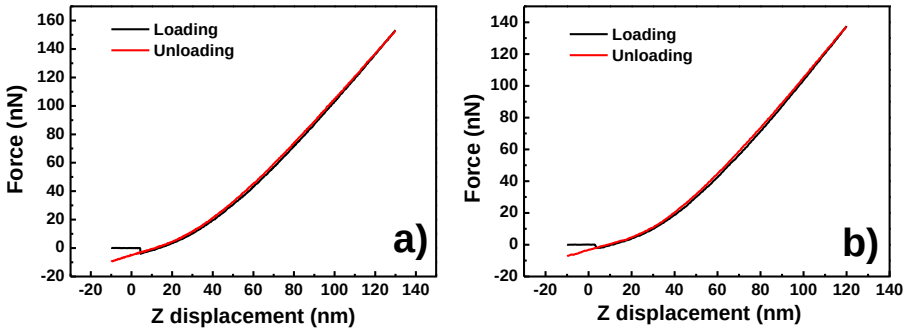



**Figure 5.5:** *Loading and unloading Force vs indentation curves acquired on a **a)**pristine and a **a)**defective graphene drums showing high reproducibility.*

tends to contract while the $SiO_2$/Si substrate shows negligible expansion [250]. This difference in behavior leads to an effective increase in the stress of the suspended graphene area that can be measured through the variation of value of $\sigma_0$ with the temperature ($\sigma_0(T)$). The 2D TEC ($\alpha_{2D}$) of graphene can be directly calculated from the dependency of $\sigma_0$ with temperature as:

$$\alpha_{2D} = -\left(\frac{\partial \epsilon}{\partial T}\right)_V = -\frac{1}{E_{2D}}\left(\frac{\partial \sigma_0}{\partial T}\right)_V, \tag{5.3}$$

where $\epsilon$ is the strain induced by temperature. This methodology was also used to estimate the TEC of defective membranes. The $F(\delta)$ curves in defective membranes displayed also excellent fitting to eq. 5.1, thus allowing also in this case an accurate determination of $\sigma_0$ and its temperature dependence, see Fig. 5.5

### 5.2.2 Experimental defect creation and characterization

In this subsection, the methodology followed by our collaborators to create and characterize defects on graphene is detailed. A controlled density of point defects was introduced in the membranes by irradiating the samples with a known dose of $Ar^+$ with incoming energy of $140eV$ in High Vacuum (HV). The irradiated area fully covers the sample holder, therefore the density of defects can be readily estimated in real time by measuring the ionic current and assuming that each $Ar^+$ removes one carbon atom [278, 279]. As reported in previous works, this technique allows the creation of controlled densities of carbon mono–

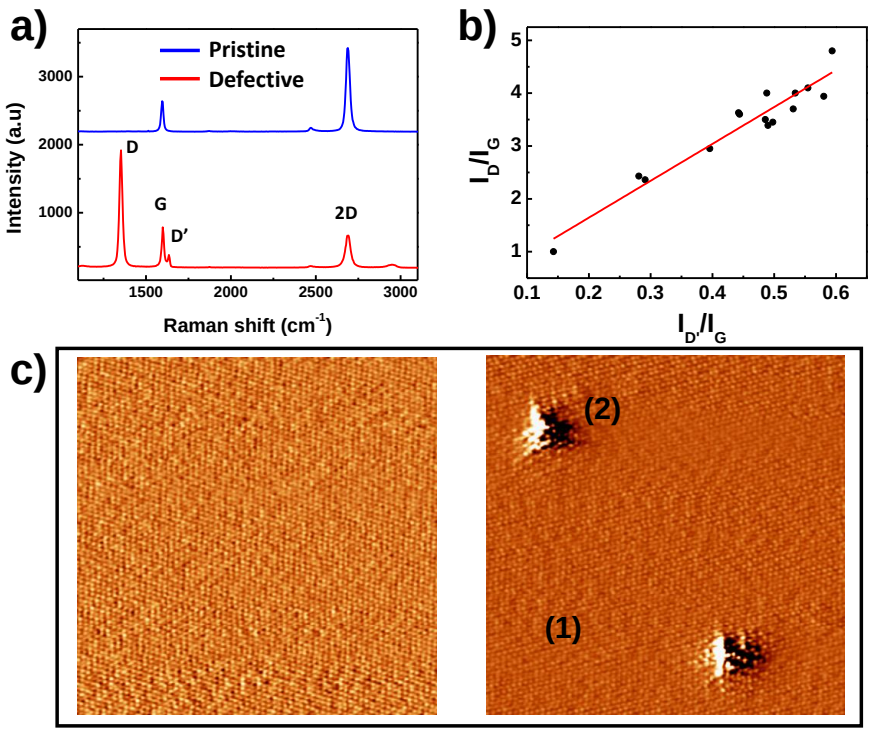



**Figure 5.6:** ***Experimental characterization of the defects.*** ***a)*** *Raman spectra of a pristine (blue) and a defective graphene membrane (red). Each peak (D, G, D' and 2D) is labeled.* ***b)*** *Values obtained for $I_D/I_G$ vs $I_{D'}/I_G$, giving $I_D/I_{D'}$=6.9± 0.5., and, thus, supporting the presence of in–plane vacancies.* ***c)*** *Atmospheric STM images of graphite before (left) and after (right) irradiation with Ar ions with the same conditions used for graphene. The size of the images is 15×15 nm$^2$. Region (2) in right panel corresponds to the threefold $\sqrt{3} \times \sqrt{3}$ perturbation due to the defect, in contrast to the hexagonal atomic periodicity typically observed by STM in pristine graphite (region 1) of the same panel.*

vacancies [4, 279]. To assure that the defects created are mono–vacancies, two different characterization techniques are followed. Firstly, Raman spectroscopy in ambient conditions over the irradiated graphene membranes is performed. The results show that the defected samples present a new peak around 1350 cm$^{-1}$, known as D peak (see Fig. 5.6a). The ratio between the intensities of the D peak and the G peak ($I_D/I_G$) informs about the density of defects induced on the sample. [280]. Moreover, additional information about the nature of defects

can be obtained from the ratio between the intensities of $D$ and $D'$ peaks of the Raman spectra as reported in reference [281]. For the case of vacancy–like defects this ratio should be about 7, and about 13 for $sp^3$ type defects. The slope of the regression line drawn in Fig. 5.6b yields an average value for this ratio of 6.9, implying the presence of in plane vacancies, instead of $sp^3$–type defects (oxidized, hydrogenated, fluorinated, etc). This analysis allows to discard the presence of chemisorbed molecules on the graphene surface.

Secondly, STM images on graphite samples before and after irradiation in the same conditions used for creating vacancies in graphene are also performed, see Fig. 5.6c. While images in graphite prior to irradiation show perfect atomic lattice, images on irradiated samples reveal small defects. These defects, visualized as small protrusions by STM, are uniformly distributed all through the sample. The threefold periodicity of the electronic perturbation observed near the defects identifies them unambiguously as point defects, i.e. smaller than lattice spacing (see region 2 in Fig. 5.6c). These STM images were obtained in air ambient conditions during 3 consecutive days without detecting any trace of image degradation by airborne molecules.

### 5.2.3 General MD Simulation Details

MD simulations were performed with the LAMMPS software suite [282] and the reactive force field AIREBO [41, 283] was used to describe the interatomic carbon interactions on the graphene sheets (see sec. 1.2.1 in chap. 1 for more details). We modelled a large pristine graphene sheet, which includes 8820 atoms and is $\sim 15 \times 15$ nm$^2$ long ($63a \times 35\sqrt{3}a$, where $a$ refers to the graphene lattice parameter, see sec. 5.1.1). The defects were included by single atom removal, i.e. they are mono–vacancies. The different concentrations and distributions of mono–vacancy defects studied are shown in Fig. 5.7. In all the simulations, we used PBC, with a cubic unit cell that fits the graphene system size in the $x$ and $y$ directions and extends 40 nm in the $z$ direction. Moreover, a time step of $\Delta t = 1$ fs and a total simulation time of 40 ns have been used in all the cases to ensure a proper sampling of the thermal fluctuations and to reach the high accuracy on the lattice parameter needed to determine changes in the TEC

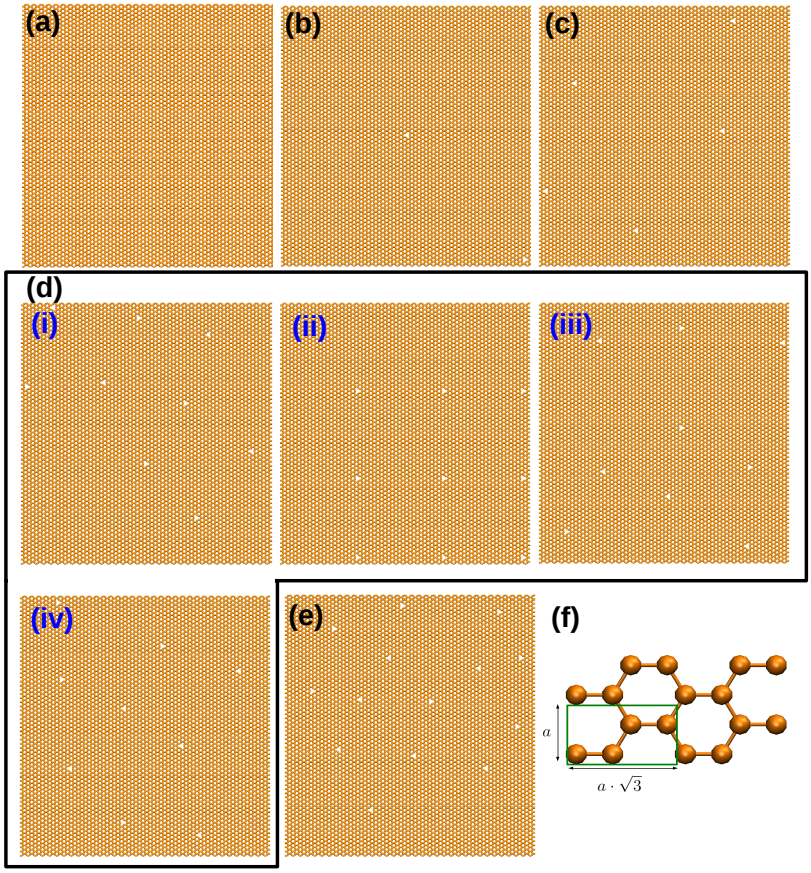



**Figure 5.7:** ***Atomic configurations for the defects densities considered in this work. a)*** *$15\times15$ nm$^2$ pristine sheet unit cell –corresponding to $63\times35$ the $a\times a\sqrt{3}$ unit cell (green rectangle) shown in* ***f)****– with 8820 carbon atoms.* ***b)-e)*** *Defective sheets, built by removing carbon atoms from the pristine membrane, and characterized by the defect concentration/density ($C_{def}$) and the average defect–defect distance ($d_{def}$):* ***b)*** *$C_{def} = 0.89\times10^{12}\ cm^{-2}$; $d_{def}$ = 10.6 nm;* ***c)*** *$C_{def} = 2.22\times10^{12}\ cm^{-2}$; $d_{def}$ = 6.70 nm;* ***d)*** *$C_{def} = 4\times10^{12}\ cm^{-2}$; $d_{def}$ = 5 nm;* ***e)*** *$C_{def} = 6.22\times10^{12}\ cm^{-2}$; $d_{def}$ = 4.06 nm. For the* ***d)*** *case, the four different defect distributions, (i) to (iv), analyzed for this concentration of defects are shown.*

and in the stress. The necessity of 40 ns long MD simulations to obtain the highly accurate value of the lattice parameter needed to distinguish changes in the TEC shall be subject of a thorough discussion in the following subsection.

### 5.2.4 NPT simulations: estimation of graphene TEC

In a first stage we aimed to estimate the TEC of graphene with and without defects using MD simulations. To that end, we analyzed the graphene dynamics at different temperatures and different defect densities. More specifically, we must study the variation of graphene size when both the temperature of the system and the defect concentration are modified. For that reason, these simulations were performed within the NPT ensemble, i.e keeping constant the temperature and pressure of the system but allowing the variation of the graphene size. The temperature and pressure of each simulation were kept constant using the Nose-Hoover thermostat and barostat. The variation of the system size was quantified via the evolution of the graphene lattice parameter for each time step, $a_t$:

$$a_t = \sqrt{\frac{lx_t \cdot ly_t}{63 \cdot 35\sqrt{3}}} \tag{5.4}$$

where $lx_t$ and $ly_t$ refer to the length of the system at each time step in the $x$ and $y$ directions, respectively. In Fig. 5.8a, we show the $a_t$ evolution in one of these simulations (pristine graphene, $T = 300\ K$). First, we analyzed the convergence of the lattice parameter during these NPT simulations by performing an incremental time average ($\langle a \rangle(t_{av})$), of its instantaneous value, $a_t$. Considering that the system is not thermally equilibrated up to $t = 20\ ns$, that incremental time average was calculated using the data of the last $20\ ns$ of simulation with the equation:

$$\langle a \rangle(t_{av}) = \frac{1}{\sqrt{N_f}} \sum_{t_i}^{t_{av}} a(t) \tag{5.5}$$

where $t_i = 20$ ns, $t_{av} = 21, \cdots, 40$ ns and $N_f = (t_{av} - t_i)/\Delta t$ refers to the number of time steps that are included in the average. The evolution of $\langle a \rangle(t_{av})$ for the pristine membrane case at $T = 300K$ is plotted in Fig. 5.8b. From that figure, we observe that $\langle a \rangle(t_{av})$ converges to a given lattice parameter value, $\bar{a}$, with an accuracy of $\sim 10^{-5}$ Å. We define that converged lattice parameter value as the average of the last ten $\langle a \rangle(t_{av})$ values, i.e. from $t_{av} = 31$ to $t_{av} = 40$ ns, see Fig. 5.8b. Repeating this procedure for different temperature and defect density values, we confirmed for all the cases the convergence of $a_t$ in $40\ ns$ of MD simulation with a precision near $10^{-5}$ Å.

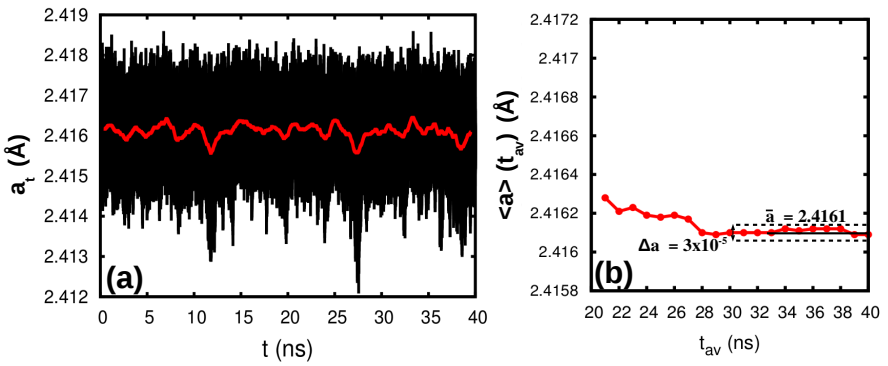



**Figure 5.8:** ***Determination of the equilibrium lattice parameter from NPT simulations.*** ***a)*** *Instantaneous value of the lattice parameter ($a_t$, in black) and running average (including the previous $10^6$ time steps, in red) for the pristine monolayer at 300 K.* ***b)*** *Incremental time average ($\langle a \rangle (t_{av})$, with red dots) calculated for the last 20 ns of the simulation shown in* ***a)****. The red line is a guide to the eye. With a black dashed line, the ten values of the $\langle a \rangle (t_{av})$ evolution used for computing the final lattice parameter value ($\bar{a}$) are delimited.*

After this analysis of the $a_t$ convergence during $40\ ns$ of NPT-MD simulation, we focus on the $\bar{a}$ dependence with the temperature. In Fig. 5.9 we plot this dependence for the pristine graphene membrane when $T = 50 - 500K$. As shown in that figure, the value of the lattice parameter decreases as the temperature increases, indicating that, as expected, the graphene shrinks when heated. The slope of this curve ($\frac{\partial \bar{a}}{\partial T}$) is directly proportional to the graphene TEC value according to the general TEC formula:

$$\alpha_{2D} = \frac{1}{a_0} \left( \frac{\partial a}{\partial T} \right)_P , \tag{5.6}$$

where $\alpha_{2D}$ refers to the 2D TEC and $a_0$ to the value of the lattice parameter at $T = 0K$. Therefore, we can compute the TEC of pristine graphene by estimating the slope of the $\bar{a}(T)$ curve in a given temperature range. Note that the variation of $\bar{a}$ when the temperature changes 100 K is about $10^{-4}$ Å, which highlights the necessity of determining the value of $\bar{a}$ with an accuracy of $10^{-5}$ Å.

We first tried to estimate that slope in the experimental temperature range,

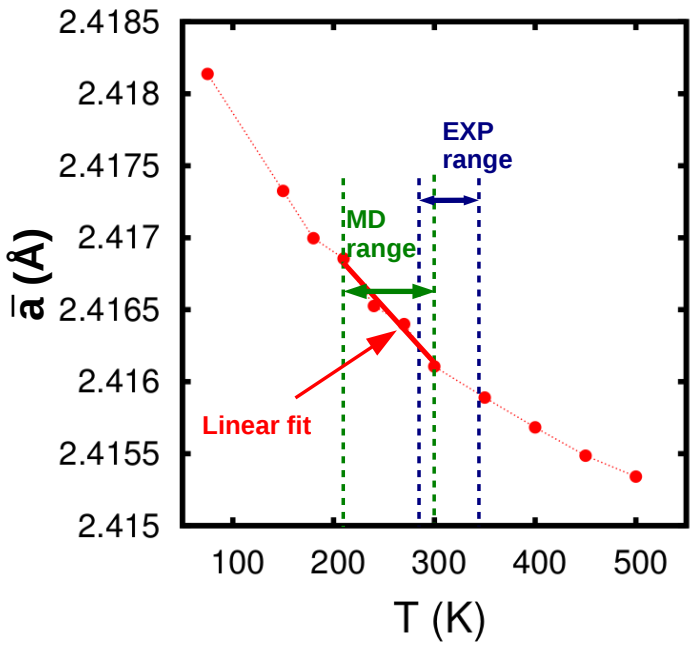



**Figure 5.9:** ***Lattice parameter as a function of temperature for pristine graphene*** *(red dots, the dashed line is a guide to the eye). The temperature ranges used in the experiments (blue) and MD simulations (green) to study the TEC evolution are shown. The lower temperatures chosen for the theoretical analysis help to achieve the precision needed to determine the changes in the lattice parameter with shorter simulation times. The methodology used for computing the TEC, i.e. a linear fit of the $\bar{a}$ data obtained from $T = 210K$ to $T = 300K$ is also here represented with a continuous red line.*

i.e from $T = 283K$ to $T = 348K$. For computing that slope, two important limitations must be considered. First, the $\bar{a}$ value must be evaluated at several temperature points within that temperature range to accurately estimate the slope of $\bar{a}(T)$. Second, in that temperature range the variation of $\bar{a}$ when the temperature changes $20\ K$ is similar to the $\bar{a}$ precision ($10^{-5}$ Å), see Fig. 5.9. Considering the length of this temperature range ($65\ K$), this second limitation could imply the ill-characterization of the $\bar{a}$ variation observed between two consecutive temperature values. Consequently, we decided to compute the graphene TEC in other temperature range, i.e. $T = 210 - 300K$, where the variation of $\bar{a}$ each $20\ K$ is larger than the $\bar{a}$ precision. We performed a linear

fit of the $\bar{a}$ data obtained at four different temperature values in that range, i.e. $T = 210, 240, 270, 300K$, see Fig. 5.9. That linear fit allows to estimate the slope of the $\bar{a}(T)$ curve, which can be then used for computing the TEC (see eq. 5.6). This procedure is repeated for estimating the graphene TEC value at different defect densities. It is worth mentioning that for defective membranes, we only analyze the $\bar{a}(T)$ dependency from $T = 210K$ to $T = 300K$, i.e. in the temperature range selected for estimating the TEC with MD simulations. We firmly believe that in the temperature range used in our MD simulations, the physical process underlaying the modification of the TEC due to the introduction of defects should be the same as the one observed in the experiments.

Additionally to the lattice parameter, other two magnitudes have been evaluated in this NPT–MD simulations to understand the role of the out–of–plane fluctuations in the TEC modification when a given mono-vacancy density is introduced in the graphene sheet. Firstly, we have computed for each time step, the standard deviation of the $z$ coordinate of the $N$ atoms of the membrane, $h_t$, as it represents the amplitude of the out–of–plane fluctuations of the membrane. The equation used for computing this magnitude is the following:

$$h_t = \frac{1}{\sqrt{N}} \left[ \sum_{i=1}^{N} (z_i(t) - z(t))^2 \right]^{1/2}; \quad z(t) = \frac{1}{N} \sum_{i=1}^{N} z_i(t) \sim 0 \tag{5.7}$$

where $z_i(t)$ refers to the instantaneous $z$ coordinate for the atom $i$. Secondly, we have analyzed the time-averaged out–of–plane fluctuation per atom, i.e. $\langle h(\bar{r}_i) \rangle$. This magnitude is computed with the expression:

$$\langle h(\bar{r}_i) \rangle = \langle (z_i(t) - z_i)^2 \rangle_t^{1/2} \tag{5.8}$$

where $z_i = \langle z_i(t) \rangle_t$ refers to the time average of $z_i(t)$. These two magnitudes are estimated for all the defective membranes here studied (see sec. 5.2.3) and for the four temperature values used for computing the TEC, i.e. $T = 210, 240, 270, 300K$

### 5.2.5 NVT simulations: understanding the origin of the TEC variation with defects

We speculated that the changes observed in the graphene thermal fluctuations with the inclusion of mono-vacancies are linked with the additional in–plane

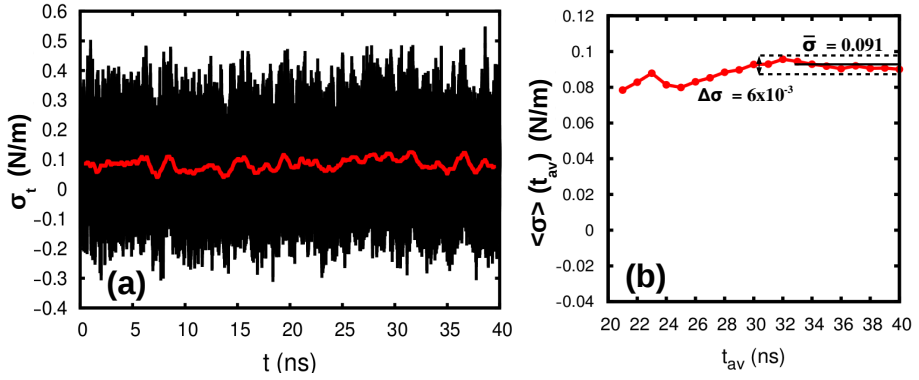



**Figure 5.10:** ***Determination of the thermal stress from NVT simulations.*** *The procedure followed to compute the thermal stress and determine its convergence follows closely the one used for the lattice parameter (see Fig. 5.8).* ***a)*** *Instantaneous value of the thermal stress ($\sigma_t$, in black) and running average (including the previous $10^6$ time steps, in red) for the pristine monolayer at 300 K.* ***b)*** *Incremental time average ($\langle\sigma\rangle(t_{av})$, red dots) for the last 20 ns of the simulation. The thermal stress final value is defined as the mean value of the last ten values of $\langle\sigma\rangle(t_{av})$, i.e from $t_{av} = 31$ to $t_{av} = 40$ ns (see dashed black line).*

stress induced in the graphene sheet by that defect creation [284]. For exploring this idea, we needed to analyze again the graphene dynamics with the temperature but allowing the change of its stress. For that reason, these new set of MD simulations was performed within the NVT ensemble. In these simulations, the system size was kept fixed. The system volume used for all the cases was the one obtained at the end of the pristine NPT simulation with $T = 210K$. The Nose-Hoover thermostat was used for keeping the temperature of the system constant. Note that the characteristics of these NVT simulations resemble the methodology used experimentally, where graphene edges are clamped.

The stress of the system at each time step ($\sigma_t$) was controlled via its relation with the pressure of the system ($P$), i.e. [46]:

$$\begin{aligned} \sigma_t &= -\frac{3V}{A} \cdot P_t \\ &= -\frac{3V}{A} \cdot [-\frac{1}{3V} \sum_{i=1}^{N} \{K^i_{xx}(t) + K^i_{yy}(t) + K^i_{zz}(t)\}], \end{aligned} \tag{5.9}$$

where $K^i_{aa} = K_{aa}(\bar{r}_i)$ are the diagonal components ($a = x, y, z$) of the per–atom stress tensor for atom $i$; $A = lx * ly$; and $V$ is the total volume of the PBC box. In Fig. 5.10a we show the $\sigma_t$ evolution for the pristine membrane at $T = 300K$. Analogously to the lattice parameter (see eq. 5.5), we also computed an incremental time average of $\sigma_t$, i.e. $\langle\sigma\rangle(t_{av})$, to analyze the convergence of this magnitude in 40 ns of NVT simulation. As shown in Fig 5.10b, the value of the system stress converges to a gives stress value,$\bar{\sigma}$, with an accuracy of $10^{-3}$ Å. We define the converged stress value analogously to the lattice parameter one, i.e as the average of the last ten $\langle\sigma\rangle(t_{av})$ values, see Fig. 5.10b. Repeating this procedure for different temperature and defect density values, we confirmed for all the cases the convergence of $\sigma_t$ in $40\ ns$ of NVT simulation.

Additionally to the stress, we also analyze another magnitude using the data extracted from these NVT simulations. This magnitude corresponds to the time average in the last 20 ns of simulation of the in–plane stress per atom $\langle K_{\parallel}(\bar{r}_i)\rangle$:

$$\langle K_{\parallel}(\bar{r}_i)\rangle = \frac{1}{2A_{at}}\langle K_{xx}(\bar{r}_i, t) + K_{yy}(\bar{r}_i, t)\rangle_t \tag{5.10}$$

where $A_{at} = A/N$ is defined as the area per atom. This magnitude and the converged value of the graphene stress ($\bar{\sigma}$) are calculated for all the defective membranes here studied (see sec. 5.2.3) and for the four temperature values used for computing the TEC, i.e. $T = 210, 240, 270, 300K$.

## 5.3 Experimental Results

Fig. 5.11a illustrates our experimentally measured $\sigma_0$ as a function of sample temperature for pristine and defective membranes. Looking to this figure, two main conclusions are extracted. First, as expected $\sigma_0$ increases with the temperature in all the cases. Secondly, the stress of defective membranes does not increase as much as that of pristine membranes with increasing temperature, i.e. they exhibit a lower |TEC|. That is confirmed in Fig. 5.11b, where we represent the value of the TEC obtained in both pristine and irradiated graphene drumheads. This result illustrates our main experimental finding: the high negative value of the TEC of graphene decreases with the introduction of vacancy-like defects, and approaches to zero ($\sim -1 \times 10^{-6}$ K$^{-1}$) for a defect

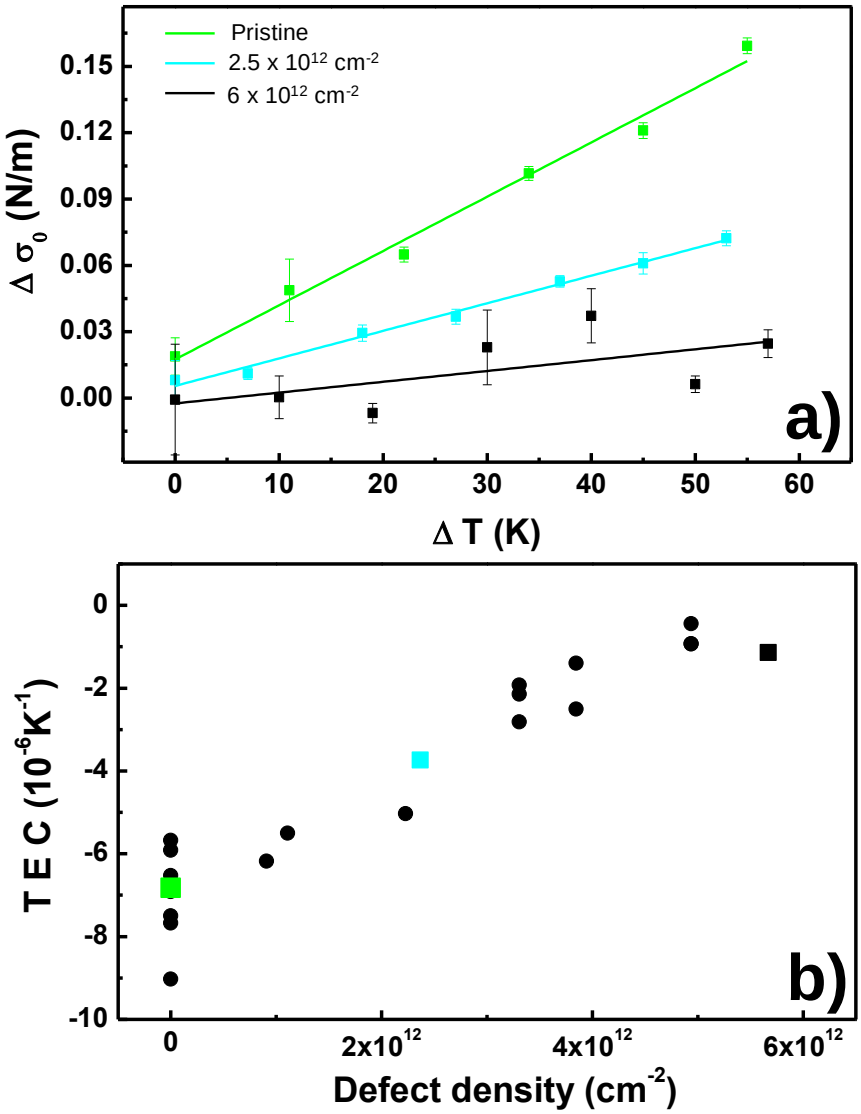



**Figure 5.11:** ***Experimental measurements of graphene TEC. a)****Plot of measured stress as a function of sample temperature. Each color corresponds to a drumhead with a different defect density. Green: pristine. Blue:* $C_{def} = 2.5 \times 10^{12}\ cm^{-2}$*. Black:* $C_{def} = 6.0 \times 10^{12}\ cm^{-2}$*.* ***b)*** *Experimentally measured TEC of different membranes as a function of the induced defect density.*

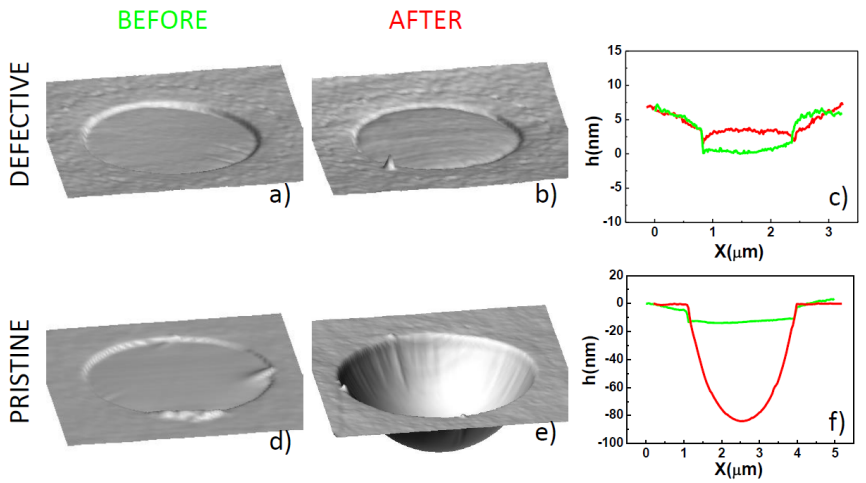



**Figure 5.12:** ***AFM images on drumheads before and after an annealing-cooling cycle from 300 to 650 K. a)*** *and* ***b)*** *panels correspond to a defective drumhead ($C_{def} = 6 \times 10^{12}$ $cm^{-2}$), while* ***d)*** *and* ***e)*** *are measured on a pristine graphene drumhead. Panels (**c**) and (**f**) are topographic profiles along the great arc of the drumheads before (green) and after (red) the annealing-cooling cycle.*

density of $6 \times 10^{12}\ cm^{-2}$ (corresponding to a mean distance between vacancies of $\sim 4$ nm). Additionally, it is worth to note that our measured mean value TEC in 7 pristine membranes, $(-7 \pm 1) \times 10^{-6}$ $K^{-1}$, agrees with previous experimental reports [248, 250](see sec. 5.1.1), which supports the validity of our technique for the characterization of thermo-mechanical properties of suspended 2D materials.

As an additional experiment to test our results through a more established approach, we performed AFM imaging of pristine and defective graphene drumheads before and after annealing-cooling cycles. Suspended pristine graphene has been reported to undergo dramatic morphological changes under such cycles [248]. These changes comprise the appearance of periodic ripples, and the bending of graphene membranes towards the substrate. Both features are caused by the thermal stress induced in the membrane by the TEC mismatch between graphene and the $SiO_2$ substrate, which has a negligible TEC. In brief, during heating graphene tends to contract in plane in order to accomplish larger out-of-plane thermal fluctuations. The in-plane strain towards the center of the

well created by this contraction makes graphene slip on the substrate towards the well. Afterwards, upon cooling, graphene expands and the area of the membrane that slid away from the $SiO_2$ during heating is maintained inside the well and buckles towards the substrate. Therefore the difference in arc length inside the well before and after a heating-cooling cycle allows estimating the TEC of graphene. Our images of pristine graphene before and after an annealing-cooling cycle (in high vacuum) from 300 to 650 K showed these morphological changes. Representative images are portrayed in Figs. 5.13d and 5.13e, where periodic ripples in the radial direction and substantial buckling towards the substrate can be observed. Based on these images, the TEC of pristine graphene can be estimated according to: $\alpha_{2D} = \frac{L_f}{L_i \Delta T}$, where $L_f$ and $L_i$ are the final and initial arc of the membrane. The value thus derived ($6.5 \times 10^{-6}$ $K^{-1}$) agrees well with that derived above and in previous reports [248]. We performed parallel AFM imaging before and after annealing-cooling cycles in membranes with a defect density of $6 \times 10^{12}$ $cm^{-2}$. In this case, we observe neither buckling nor periodic ripples (Figs. 5.13a and 5.13b), indicating that the TEC mismatch between the substrate and the graphene membrane approaches to zero i.e. a strongly reduced |TEC| in defective graphene, further supporting the results presented above.

## 5.4 MD-NPT Simulation Results

### 5.4.1 Theroretical graphene TEC Estimation

As a first step to understand our measurements, we performed constant pressure NPT-MD simulations to determine the lattice parameter versus temperature for pristine and defective graphene, see sec. 5.2.4 for more details. Fig. 5.13a displays the lattice parameter results obtained in our NPT-MD simulations for the four temperature values selected for estimating the TEC (T=210, 240, 270, 300K, see sec. 5.2.4) and three different graphene membranes: pristine (green circles), $C_{def} = 2.22 \times 10^{12}$ $cm^{-2}$ (purple circles) and $C_{def} = 4 \times 10^{12}$ $cm^{-2}$ (black circles). By performing a linear fit of this $\bar{a}(T)$ data, we can analyze the $\bar{a}$ variation with the temperature for the pristine membrane and how it changes

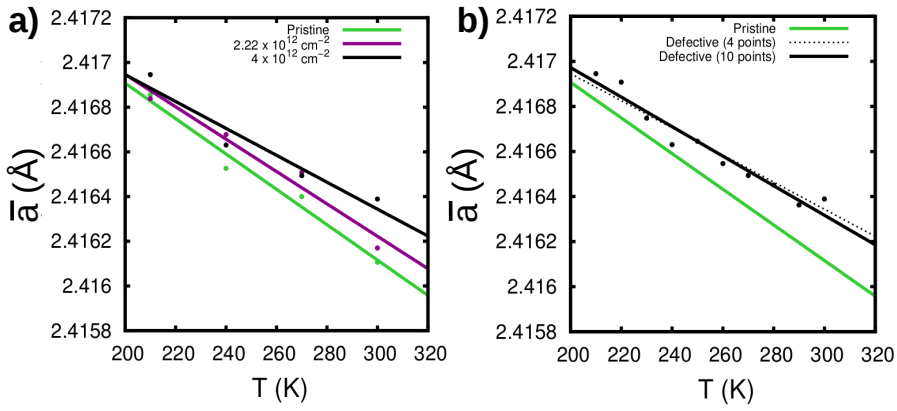
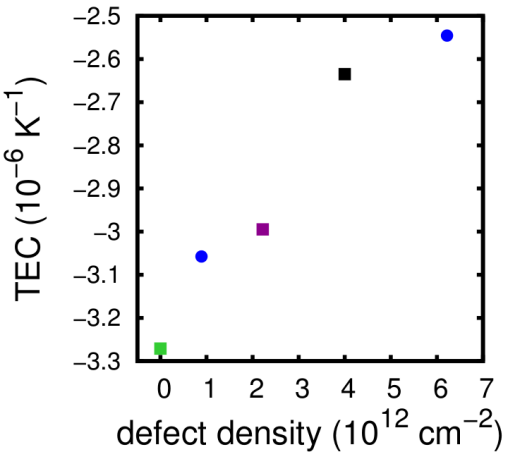



**Figure 5.13:** ***Theoretical determination of the*** $\bar{a}(T)$ ***dependence from NPT simulations.*** ***a)*** *Lattice parameter versus temperature for pristine (green) and two defective graphene membranes:* $C_{def} = 2.2 \times 10^{12}$ *cm*$^{-2}$ *and* $C_{def} = 4 \times 10^{12}$ *cm*$^{-2}$ *(configuration i). The TEC is determined by the slope of the corresponding linear fits (solid lines).* ***b)*** *Lattice parameter versus* $T$ *for one of the defective cases (*$C_{def} = 4{\times}10^{12}$ *cm*$^{-2}$*, configuration i) and the corresponding fit using 10 different temperatures (solid black line). The fits for the pristine case (solid green line), and the same defect configuration using 4 different temperatures (dotted black line) are also shown for comparison.*

with the introduction of mono-vacancies. The linear fits obtained for these three graphene membranes (pristine, $C_{def} = 2.22{\times}10^{12}\ cm^{-2}$,$C_{def} = 4{\times}10^{12}\ cm^{-2}$) are also shown in Fig. 5.13a. From that figure, it is clear that the decrease of $\bar{a}$ with the temperature in defective membranes is smaller than in pristine graphene. This result is robust with respect to the number of sampling temperatures used for performing the linear fit, as we continue observing that difference when it increases from four to ten, see Fig. 5.13b. That supports our choice of sampling only four temperatures for estimating the modification of the pristine graphene $\bar{a}(T)$ tendency with the creation of defects.

We can then quantify the TEC using the standard TEC formula, $\alpha_{2D} = \frac{1}{a_0}\left(\frac{\partial \bar{a}}{\partial T}\right)_P$ (see more details in sec. 5.2.4). In Fig 5.14 we plot the TEC value obtained for all the graphene membranes analyzed in this work, see Fig. 5.7. These values are also detailed in the table below. Note that in the $C_{def} =$




| $C_{def}$ ($10^{12} cm^{-2}$) | 0 | 0.89 | 2.22 | 4 (cnf. i) | 4 (cnf. ii) | 4 (cnf. iii) | 4 (cnf. iv) | 6.22 |
|---|---|---|---|---|---|---|---|---|
| $\bar{a}_{T=210K}$ (Å) | 2.41686 | 2.41684 | 2.41684 | 2.41695 | 2.41686 | 2.41690 | 2.41696 | 2.41699 |
| $\bar{a}_{T=240K}$ (Å) | 2.41653 | 2.41662 | 2.41668 | 2.41663 | 2.41670 | 2.41666 | 2.41662 | 2.41676 |
| $\bar{a}_{T=270K}$ (Å) | 2.41640 | 2.41632 | 2.41651 | 2.41649 | 2.41645 | 2.41651 | 2.41653 | 2.41660 |
| $\bar{a}_{T=300K}$ (Å) | 2.41611 | 2.41620 | 2.41617 | 2.41639 | 2.41630 | 2.41631 | 2.41633 | 2.41643 |
| **m** ($10^{-6}$Å$K^{-1}$) | -7.91342 | -7.39712 | -7.24532 | -6.02112 | -6.45572 | -6.4118 | -6.61172 | -6.16047 |
| $a_0$ (Å) | 2.41922 | 2.41923 | 2.41936 | 2.41938 | 2.41938 | 2.41938 | 2.41938 | 2.41995 |
| **TEC** ($10^{-6} K^{-1}$) | -3.27106 | -3.05763 | -2.99473 | -2.4887 | -2.66834 | -2.65018 | -2.73282 | -2.5457 |

**Figure 5.14:** ***Theoretical determination of the TEC from NPT simulations.*** ***(Top)****TEC as a function of the defect density. Data for* $C_{def} = 4 \times 10^{12}$ $cm^{-2}$ *is an average of the four different defect configurations (see Fig. 5.7 in sec. 5.2.3).* ***(Bottom)****Lattice parameter* $\bar{a}$ *for each temperature and defect concentration, slope of the* $\bar{a}(T)$ *curve* ***(m)****, lattice parameter value at T = 0 K* ($\boldsymbol{a_0}$)*, and calculated TEC values.*

$4 \times 10^{12}$ cm$^{-2}$ case, we have computed the TEC value for four different defects configurations, see table in Fig 5.14 and Fig. 5.7. The data shown in the figure (black square) corresponds to the average of these four TEC values.

From Fig. 5.14, we extract two main results. Firstly, our TEC for pristine graphene, $(-3.27 \times 10^{-6}\ K^{-1})$, is in excellent agreement with previous *ab initio* [256] and classical–potential [46, 257] theoretical studies. Secondly, our results for the defective systems show that the introduction of defects leads to a reduction of the |TEC|, in agreement with our experimental observations. This result is robust with respect to the defect distribution, as similar TEC values are

obtained for all the defect configuration studied when $C_{def} = 4 \times 10^{12}$ m$^{-2}$, see table in Fig 5.14.

Comparing the experimental and simulated TEC trends, see Fig 5.14 and Fig. 5.11 respectively, we observe that both are very similar. The |TEC| decreases rapidly as the concentration of defect increases, and tends to saturate for concentrations beyond $4 \times 10^{12}$ cm$^{-2}$. In spite of this remarkable agreement between experimental and simulated trends, there are quantitative discrepancies regarding the TEC value for pristine graphene, as our theoretical results predict a TEC almost two times smaller than our experiments. This discrepancy has been already pointed out in previous experimental works [248, 250], see sec. 5.1.1. According to recent MD simulations [46], |TEC| increases with system size for square $L \times L$ unit cells in the range $L \sim 10 - 100$ nm. This behavior can be understood in terms of the effective cutoff introduced by the finite system size in the long-wavelength out–of–plane vibration modes that are responsible for the negative TEC of graphene. This interpretation implies that the smaller difference observed between the TEC experimental and theoretical values in defective systems ($TEC_4^{exp} \sim -2.4 \times 10^{-6}$ K$^{-1}$ and $TEC_4^{theo} \sim -2.63 \times 10^{-6}$ K$^{-1}$, see Figs 5.14 and 5.11) arises from the reduction of the average out–of–plane fluctuation in that graphene membranes. In the next subsection, we inspect that possibility by analyzing the changes induced in the out–of–plane fluctuations by the combined effect of temperature and the presence of defects.

### 5.4.2 Influence of defects in the graphene out-of-plane fluctuations

In order to understand the microscopic mechanism behind the TEC reduction, we explore the changes induced in the out–of–plane fluctuations by the combined effect of temperature and the presence of defects. In Fig. 5.15a-b, we show the evolution of the spatially averaged amplitude of the graphene out–of–plane oscillations, i.e. $h(t)$ (see definition in sec. 5.2.4) for the pristine membrane at $T = 210K$ and $T = 300K$. Considering the effect of these out–of–plane fluctuations in the graphene size, i.e. the larger they are, the smaller is the value of the lattice parameter (see Fig. 5.15c), the evolution of this magnitude must be anti-correlated with the evolution of the lattice parameter. That anti-correlation

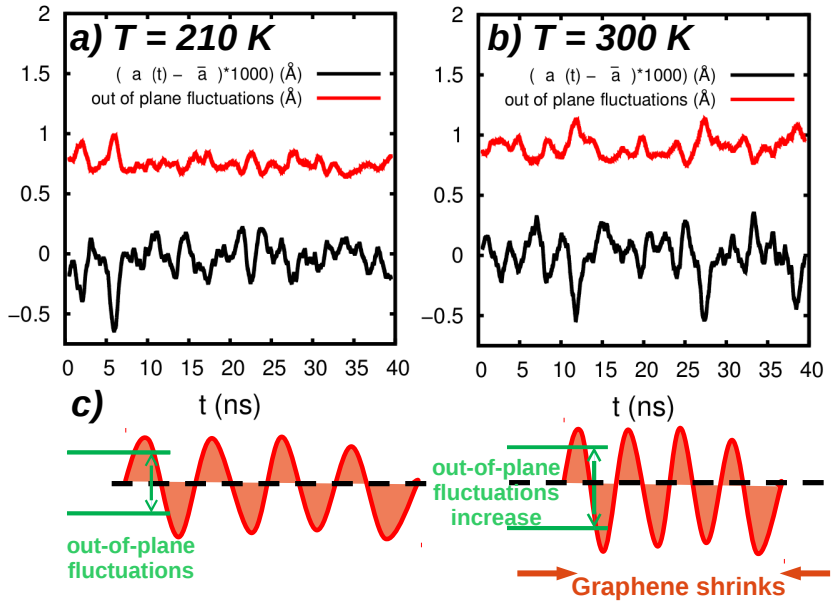



**Figure 5.15:** ***Evolution of the average out–of–plane fluctuation during NPT simulations.*** *Time evolution of the spatially–averaged amplitude of the out–of–plane fluctuations, $h(t)$ (red, see sec. 5.2.4) for the graphene pristine membrane at two different temperature values:* ***a)*** $T = 210$ *K and* ***b)*** $T = 300$ *K. The fluctuations of its lattice parameter for these two temperature values are also included (black).* ***c)****Schematic representation of the relation between the out–of–plane fluctuations and the graphene size.*

is observed in our simulations, see Figs. 5.15a-b. For the two temperature values here shown, we observe that whenever the amplitude of the fluctuations increases (in black) the lattice parameter decreases (in red). This result shows that our NPT-MD simulations are able to capture the anti-correlation between these two quantities which seem to be key in understanding the origin of TEC of graphene without and with defects. Additionally, it must be noted that these simulations also capture the increase of graphene out–of–plane fluctuations with the temperature, as $h(t)$ is larger when $T = 300K$ than for $T = 210K$. This reinforces the validity of these simulations for estimating the graphene negative TEC, as it arises from the increase with the temperature of the out–of–plane fluctuations, see sec. 5.1.1.

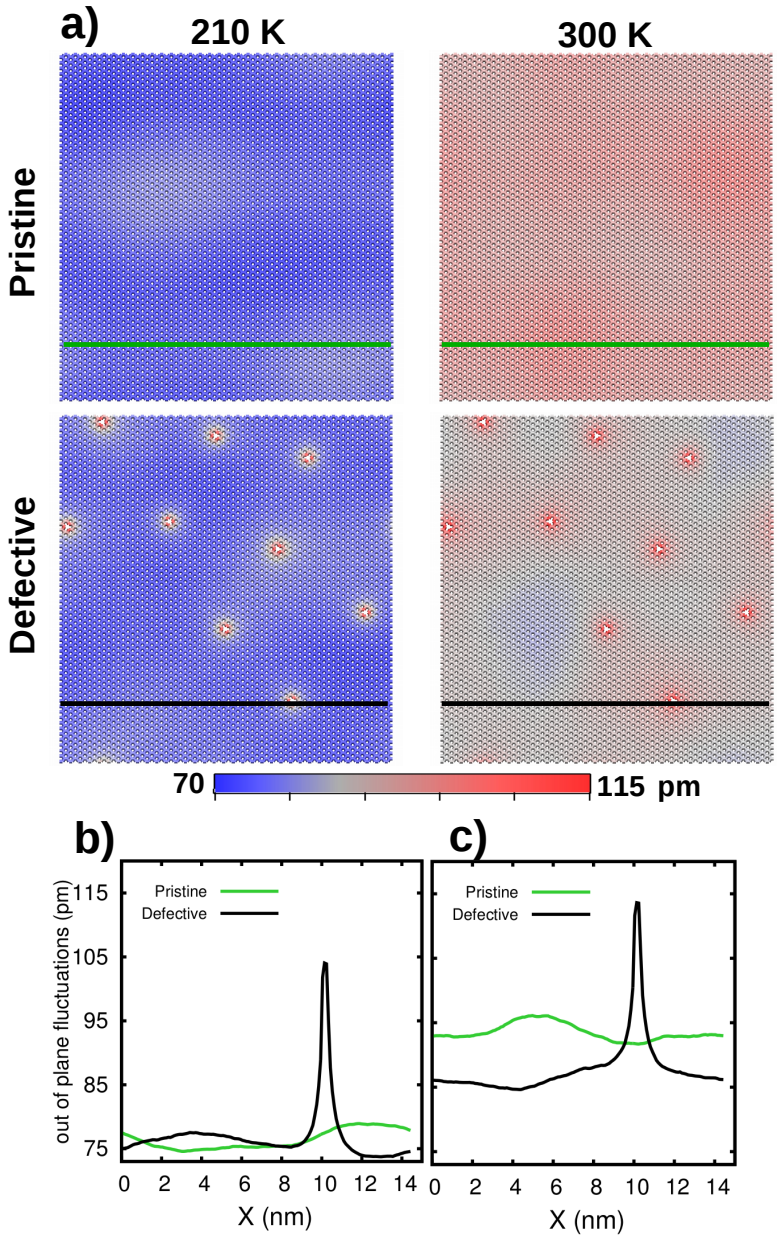



**Figure 5.16:** ***Analysis of the amplitude of the out–of–plane fluctuation per atom obtained in our NPT simulations. a)*** *Amplitude of the out–of–plane fluctuation per atom, $\langle h(\bar{r}_i)\rangle$, (average of the $z$–coordinate of atom $i$ over the last 20 ns) for pristine and defective graphene at two temperatures ($T = 210, 300$ K).* ***b)*** *Profile of $\langle h(\bar{r}_i)\rangle$ along the green (black) line in panel* ***a)*** *for pristine (defective) graphene at $T = 210$ K.* ***c)*** *Same as b) for $T = 300$ K.*

Further insights can be gained by looking at the real space distribution (atom by atom) of the time averaged amplitude of the out–of–plane fluctuations , i.e. $\langle h(\bar{r}_i)\rangle$, with $\bar{r}_i$ defined as the $xy$ in–plane position of atom $i$, see sec. 5.2.4. That magnitude is displayed in Fig. 5.16a for two different temperatures and in the presence of defects. These 2D maps capture two important features: First, for pristine graphene, the dominant red color at $T = 300\ K$, in contrast to the blue color at $T = 210\ K$, reflects that temperature is responsible for the increase in the amplitude of out–of–plane fluctuations for each atom, leading to an actual in–plane contraction of the material. Second, defective graphene exhibits overall smaller fluctuations than pristine graphene. Although the atoms that are first neighbors of the defects have larger oscillation amplitudes, fluctuations have been drastically reduced on the rest of the graphene sheet. This effect is specially seen at $T = 300\ K$, where the color change between pristine and defective membranes is larger. That difference can be quantified using the line profiles shown in Fig. 5.16b–c. Summarizing, our simulations confirm that the TEC reduction in the presence of defects is induced by the reduction of out–of–plane fluctuations, as was previously suggested (see sec. 5.1.2).

## 5.5 MD-NVT Simulation Results

At this stage, we speculate that the quenching of out–of–plane fluctuations is linked with an additional in–plane stress induced in the graphene sheet by the defects. The presence of an uniform tensile strain is known to reduce the out–of–plane fluctuations [284]. We explored this idea with NVT-MD simulations, see sec. 5.2.5 for more details. Fig. 5.17 shows the converged thermal stress ($\bar{\sigma}$) for pristine and defective graphene membranes as a function of temperature. In this figure we note that the thermal stress of pristine graphene increases much faster than the defective one, as was observed in the experiments (see Fig. 5.11a). This behavior allows us to independently confirm that a decrease on the thermal strain indeed translates into a reduction of the TEC.

In our simulations, we are able to decompose the total stress into a sum of local contributions for each of the carbon atoms, see sec. 5.2.5. Fig. 5.18a presents the time average of the sum of the in–plane diagonal components of the

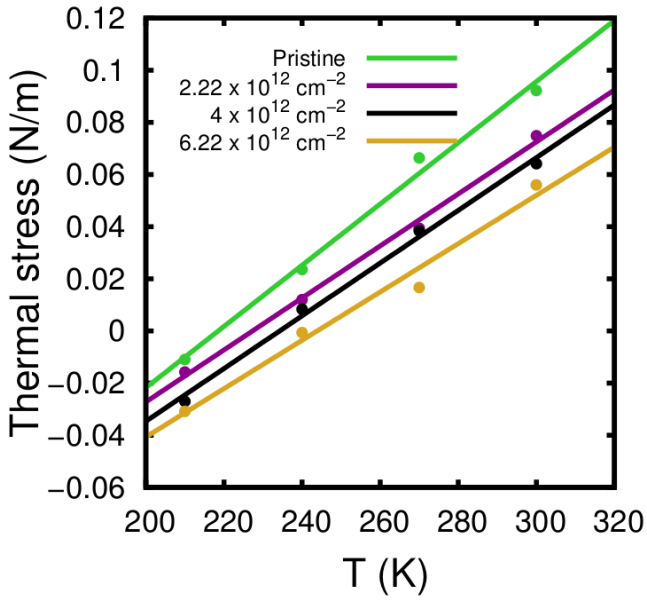



**Figure 5.17:** ***Thermal stress as a function of temperature for different defect concentrations.*** *Thermal stress versus $T$ for each defect concentration (bullets) and the corresponding linear fits (lines). The result for $C_{def} = 4 \times 10^{12}\ cm^{-2}$ corresponds to the (i) defect configuration, see Fig. 5.7.*

$i$–atom stress tensor $\langle K_{\parallel}(\bar{r}_i)\rangle = \langle (K_{xx}(\bar{r}_i) + K_{yy}(\bar{r}_i))/2A_{at}\rangle_t$ as a 2D stress map for both pristine and defective graphene at two different temperatures. In–plane compressive (positive) stress is represented in red, while blue corresponds to in–plane tensile (negative) stress, that would tend to expand locally the graphene sheet. For pristine graphene at $T = 210\ K$, we have an even distribution of very small regions with low positive and negative stress that cancel out. This result is consistent with the fact that the simulation box size is fixed to the equilibrium value of pristine graphene at $T = 210\ K$. When temperature is raised to 300 K, the 2D stress map turns red, revealing an homogeneous increase of the in–plane compressive stress in the system (see Fig. 5.18b–c) as observed experimentally. This means that, if fixed boundary conditions were relaxed, graphene would shrink in–plane.

The 2D stress map for defective graphene at $T = 210\ K$ (see Fig. 5.18a)

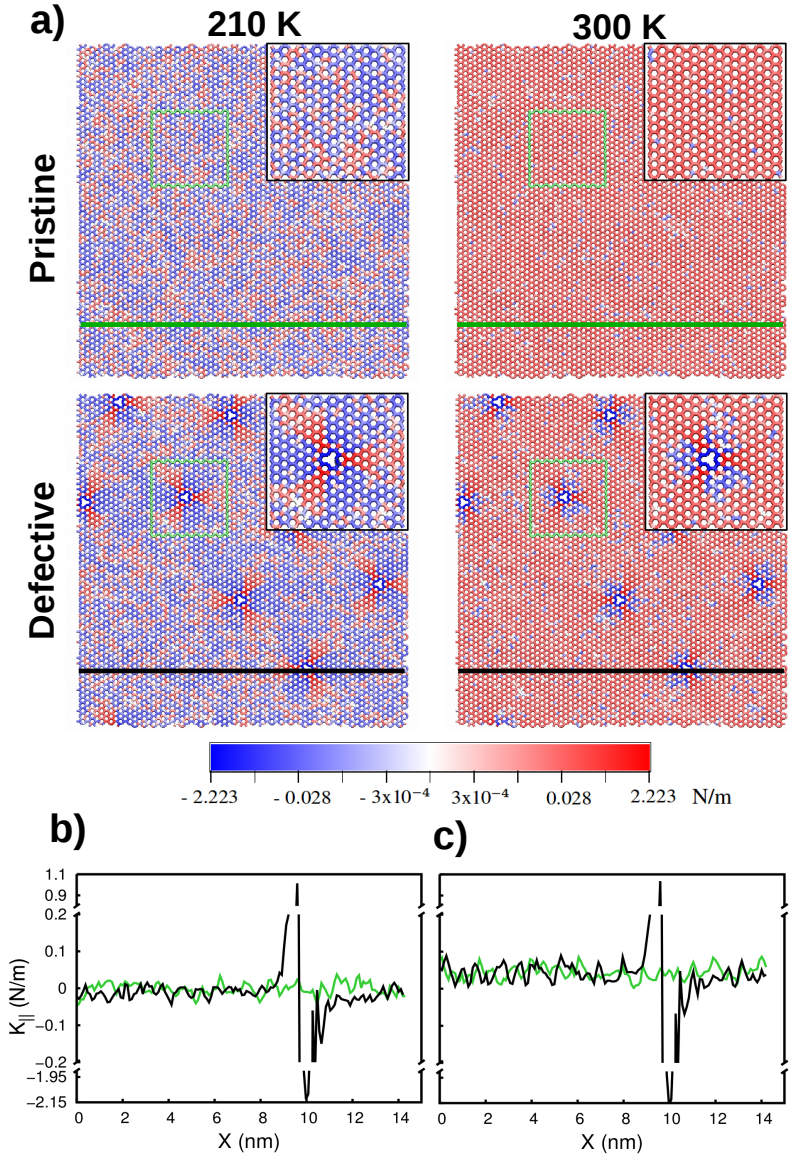



**Figure 5.18:** ***2D stress maps obtained from our NVT simulations. a)****Sum of the in–plane diagonal components of the $i$–atom stress tensor, $\langle K_{\parallel}(\bar{r}_i)\rangle$ , (normalized by the surface area per atom $A_{at} = 0.02529$ nm$^2$) for pristine and defective graphene at $T = 210, 300$ K. The data here shown corresponds to a time average of the $\langle K_{\parallel}(\bar{r}_i)\rangle$ value in the 20 ns of simulation.* ***b)****Profile for $\langle K_{\parallel}(\bar{r}_i)\rangle$ along the green (black) line in panel a) for pristine (defective) graphene at $T = 210$ K.* ***c)****Same as b) for $T = 300$ K. Notice the logarithmic scale used in panels b)–c).*

looks similar to the one of the pristine case in areas far from the defects (e.g., the left bottom corner). However, around defects, we observe a three-fold symmetry stress distribution, where areas of tensile (blue) and compressive (red) stress alternate. Overall, the map shows a predominance of blue areas with tensile stress. If fixed boundary conditions are released, this stress would favor an in-plane expansion that naturally leads to a larger equilibrium lattice parameter for the defective case, in agreement with our NPT results (Fig. 5.13a). At $T = 300K$, we observe an increase of the in–plane compressive stress similar to the one found in the pristine case. However, at variance with the latter, small patches of tensile stress can still be found around the defects. These patches are responsible for a reduction of the global compressive stress in the system, that would lead, if boundary conditions are removed, to a smaller in–plane contraction compared to the pristine case.

The maximum tensile stresses observed around defects for both temperatures (see Figs 5.18b and 5.18c) are $\sim 2.15$ N/m, that corresponds to a local strain of 0.65 %. Theoretical calculations based on the self-consistent screening approximation and atomistic Monte-Carlo simulations [284], predict a very strong attenuation of the out–of–plane thermal fluctuations in stiff membranes, as graphene, with the presence of rather weak uniform strains ($\sim 0.4$ %). Strain fields less than 1% are enough to suppress all the anharmonic effects. This strong dependence with the strain provides the fundamental link between our results from the NPT and NVT simulations. Defects support short–wavelength normal vibrations, but the local strain that they induce contributes to the quenching of the long–wavelength thermal fluctuations that are mainly responsible for the in–plane graphene contraction. The remaining fluctuations lead to the reduced negative TEC measured in our experiments and simulations in the presence of defects. Thus, the stress distribution induced by the defects, in particular, the presence of areas of tensile stress, is responsible for the |TEC| reduction.

## 5.6 Conclusions

In summary, we measured the TEC of graphene using a novel approach that can be extended to other 2D materials. We also shown that the TEC can be

tuned through the controlled introduction of vacancy-like defects, reducing it to almost zero for a modest defect concentration (corresponding to a mean distance between vacancies of $\sim 4$ nm). Our long MD simulations reproduce this trend, validate the experimental approach, and provide a comprehensive understanding of this phenomenon in terms of the quenching of out–of–plane fluctuations induced by the patches of tensile strain introduced by the defects. Although textbook physics of pristine graphene has attracted much attention, our work proves that the unavoidable presence of small defect densities in real–life graphene has profound implications. Our insight provides a guide for future proposals of either electronic or thermal defect engineering [285]. Along this line, we suggest controlled creation of low defect densities to reduce the TEC mismatch between graphene and substrates that cause undesired strains in devices. Our results have also to be considered in the interpretation of previous reports on the interplay of defects and mechanical properties. In particular, they constitute the first evidence supporting the conjecture of graphene stiffening by quenching of long–wavelength phonons by monovacancies [4], paving the way for the tailoring of the mechanical properties of graphene and other 2D materials where out–of –plane thermal fluctuations are relevant.

# 6 General conclusions

Nanoscale systems have shown, in last decades, novel and different properties compared with those of their macroscopic counterparts. This has encouraged the scientific community to try to understand and control the properties of these nano-systems to promote their integration in different nanotechnological applications. In this context, the work presented in this thesis was devoted to study the dynamical properties of different nano-systems and to understand how they can be drastically modified by introducing atomic-scale modifications on both the environment and the intrinsic structure of the nano-system. By combining the use of "all-atom" Molecular dynamics simulations with "state-of-the-art" computing architectures, such as GPUs, we have been able to describe the atomistic mechanisms that influence the dynamical properties of a large variety of nanoscale systems, ranging from 2D materials over ssDNAs assemblies to proteins. The results and methodology here developed pave the way for a tight control of the properties at the nanoscale.

In chapter 2, we have inspected the adsorption process of the protein Immunoglobulin G (IgG) on graphene with both explicit (all the water atoms are explicitly included on the system), and implicit (the solvent is described as a continuous dielectric medium) solvation methods. In the former case, we have not observed major changes in the IgG secondary structure upon adsorption (less than a $5\%$). However, implicit solvation results have shown a major structural rearrangement of the protein upon adsorption, which is not compatible with the experimentally demonstrated preservation of the IgG bio-functionality on graphene [56]. By performing a detailed energy decomposition analysis of both simulations, we have been able to unveil the mechanisms behind the IgG

adsorption over graphene and the importance of the atomistic description of the solvent in those mechanisms. Regarding the adsorption mechanisms, our explicit solvent simulation analysis has evidenced that the protein-surface attraction is counterbalanced by the cost of breaking both the surface and protein solvation shells, and thus the IgG adsorption proceeds as a consequence of a subtle effect, i.e. the entropic water thermal oscillations. This energetic balance is not captured using implicit solvent, as in this case the cost of breaking the solvation shells is underestimated. Moreover, the gain of water-water electrostatic energy during adsorption is also not considered in implicit solvent simulations. These two factors prompt that the IgG adsorbs instantaneously over graphene and with a pronounced loss of solute-solvent interaction, resulting in large electrostatic unbalances inside the protein, which ultimately leads to its unfolding.

Chapter 3 presents how the hydration influences the mechanical response of a self-assembled monolayer (SAM) of ssDNAs. Our MD simulation results at eleven hydration levels (from H=0 to H=600 water per ssDNA) have shown that the mechanical response of this system, here quantified via its curvature, goes through three well differentiated stages: (1) at low hydration levels, the curvature increases with the hydration; (2) for H>100 water/ssDNA the curvature decreases until (3) it saturates for H$\leq$400 water/ssDNA. These three stages are compatible with previous experiments, which analysed the surface stress variation of an analogous SAM with the relative humidity of the system. By performing a detailed energetic and structural analysis, we have been able to put in contact the atomistic water-ssDNA interactions with the collective SAM structural and mechanical response. At low hydration levels (1), the high water-ssDNA affinity prompts the ssDNAs to become stiffer with the hydration and to approach to each other, leading to a larger upward deflection of the SAM as the hydration increases. For 100<H<400 water/ssDNA (2), the rise of the water-water affinity prompts an enlargement of the SAM nanopores with the hydration, i.e the ssDNAs stretch with the hydration, which ultimately hinders the increase of the curvature. For H>400 water/ssDNA (3), the complete filling of the nanopores with water molecules prompts the new water molecules added to the system to remain outside the SAM and, thus, the SAM structural and

mechanical properties barely change. We have completed this study by analyzing how this hydration-driven mechanical response changes when the SAM grafting density is $\sim$ $1.3$ times smaller. Our results confirm that the reduction of the grafting density introduces a new ssDNAs structural tendency at very low hydration levels as a consequence of the larger free space existent in the nanopores, which influences the SAM mechanical response. All in all, the results presented in this chapter unveil the surprising active role of the water molecules in the structure and mechanical properties of these SAMs.

In chapter 4, we have explored the effect of single amino-acid mutations on the structure and dynamics of the Azurin protein. This study was performed not only to characterize the effect of these atomic-scale mutations on the dynamics of a protein, but also to understand the change of the electronic transport process through a gold-Azurin-gold junction observed experimentally by replacing one of its amino-acids with a cysteine. By simulating the unrestrained dynamics in water of the mutants studied in the experiments, we have shown that the introduction of mutations does not practically alter the Azurin structure but it reduces the Azurin global fluctuations up to a $8\%$, i.e. the Azurin becomes stiffer. To analyze the influence of this mutation-induced stiffening effect on the Azurin-gold interaction, we have also simulated the free adsorption process of both "wild-type" and mutant Azurin variants over a gold surface. Four different initial Azurin-surface orientations were tested to obtain a wider statistics of the Azurin adsorption process. Our adsorption simulations have revealed a high mobility for the wild-type protein during adsorption, as it was able to undergo major reorientation maneuvers to adsorb along one of the two preferred configurations: lying-down with the cysteines in contact with the surface or through the hydrophobic patch. In contrast, the mutant has shown a smaller reorienting ability as it always adsorbs along its initial Azurin-gold orientation. These contrasting results suggest that quenching the Azurin fluctations via mutation creation may significantly alter its adsorption process over gold, even changing the Azurin-gold configuration. This result paves the way for an understanding of the experiments, as variations on the Azurin-gold configuration could ultimately affect the Azurin conductivity on the junction.

In chapter 5, we have examined the role of atomic-scale modifications on the thermo–mechanical properties of graphene. This work combines both experiments and simulations as it has been performed in close collaboration with a experimental group of our University. Our aim was to unveil the influence of a dilute density of monovacancies on the out-of-plane thermal fluctuations of graphene, which cause the graphene to contract when heated (negative thermal expansion coefficient). Considering this relation between the graphene thermal expansion coefficient (TEC) and its out–of–plane fluctuations, we have decided to estimate this quantity both experimentally and with MD simulations for different monovacancy densities (from 0 to 6 $\times 10^{12}$ cm$^{-2}$). The experiments have demonstrated that the graphene TEC can be tuned through the controlled introduction of monovacancies, reducing it to almost zero when the mean distance between vacancies is $\sim$ 4 nm. By performing 40ns-long NPT-MD simulations, we have been able to quantify variations in the graphene lattice parameter with a precision of $10^{-5}$Å, which allows us to accurately estimate the negative graphene TEC and confirm that it reduces with the introduction of mono-vacancies. To understand this reduction, we have analyzed the stress distribution along the different graphene sheets via 40ns-long NVT-MD simulations, concluding that monovacancies create strain fields in their surrounding areas, which direclty affect the out–of–plane graphene modes.

In summary, this work presents a comprehensive study of the role of atomistic interactions on different dynamical properties at the nanoscale. In particular, we have revealed the solvent-mediated mechanisms occurring in the atomistic scale during protein adsorption on a hydrophobic surface. Furthermore, we have unveiled that the water atomic interaction plays an active role in the structure and mechanical response of complex biological systems such as ssDNAs assemblies. Finally, we have demonstrated that the introduction of atomic-scale modifications on the intrinsic structure of two different nano-systems (protein, graphene) significant changes their mechanical properties, explaining also the physical mechanisms behind such impressive results.

# Conclusiones generales

Los sistemas en la nanoescala o "nano-sistemas", han demostrado en las últimas décadas tener propiedades diversas y novedosas comparadas con las de sus análogos en la macroescala. Esto ha incentivado el interés de la comunidad científica por comprender y modificar las propiedades de estos nano-sistemas para favorecer su integración en aplicaciones nanotecnológicas. En este contexto, el trabajo que presentamos en esta tesis está dedicado a estudiar las propiedades dinámicas de diferentes nano-sistemas y cómo pueden ser modificadas drásticamiente por medio de modificaciones a nivel atómico, tanto en su estructura intrínseca como en su entorno. Combinando el uso de simulaciones de Dinámica Molecular (MD) "all-atom" con el uso de nuevas arquitecturas computacionales, como las tarjetas gráficas, hemos logrado describir los mecanismos atómicos que influyen en las propiedades dinámicas de una gran variedad de sistemas en la nanoescala, desde materiales 2D hasta proteínas, pasando por monocapas de hebras de ADN. Los resultados y metodología desarrollados en esta tesis sientan las bases necesarias para el control preciso de las propiedades en la nanoescala.

En el segundo capítulo, hemos inspeccionado el proceso de adsorción de la proteína Immunoglobulina G (IgG) sobre grafeno, tanto con métodos de solvatación explícitos (todos los átomos se incluyen explícitamente en el sistema) como implícitos (el solvente se describe como un medio continuo dieléctrico). En el primer caso, no se observaron grandes cambios en la estructura secundaria de la IgG tras la adsorción (menos de un $5\,\%$). Sin embargo, los resultados de la solvatación implícita han mostrado un mayor cambio en la estructura terciaria de la proteína tras la adsorción que se contradice con la preservación de la bio-funcionalidad de la IgG sobre el grafeno demostrada experimentalmente.

Por medio de un análisis detallado de las diferentes contribuciones energéticas en ambos tipos de simulaciones, hemos conseguido descrifrar los mecanismos que gobiernan la adsorción de la IgG sobre el grafeno y la importancia de la descripción atomística del solvente en los mismos. En relación a los mecanismos de adsorción, nuestras simulaciones de solvatación explicita han demostrado que la atracción superficie-proteína se contrarresta con el coste energético de ruptura de las capas del solvatación que rodean tanto a la proteína como a la superficie. Como consecuencia, la adsorción de la proteina es un proceso lento, el cual ocurre gracias a las fluctuaciones térmicas/entrópicas del agua. Este balance energético no es capturado por los métodos de solvatación implicita, pues en este caso se subestima el coste energético de ruptura de las capas de solvatación. Además, en las simulaciones de solvente implícito no se tiene en cuenta la ganancia de energía electrostática interna del agua (se forman mas puentes de hidrógeno agua-agua) durante la adsorción. Estos dos factores llevan a que la IgG se adsorba de manera instantánea sobre el grafeno, perdiendo significativamente interacciones electrostáticas con el solvente, lo que causa un gran desquilibrio electrostático en el interior de la proteína y por tanto la pérdida de su estructura terciaria.

El tercer capítulo versa sobre la influencia de la hidratación en la respuesta mecánica de una monocapa autoensamblada de hebras simples de ADN (ssDNA). Hemos realizado once simulaciones de Dinámica Molecular usando en cada una de ellas un diferente valor de la hidratación (desde H=0 hasta H=600 moléculas de agua por cada ssDNA).Los resultados muestran que la respuesta mecánica de este sistema, cuantificada aquí mediante su curvatura, pasa por tres estados claramente diferenciados: (1) en niveles bajos de hidratación, la curvatura aumenta con la hidratación; (2) para niveles H>100 agua/ssDNA la curvatura disminuye hasta su saturación en (3) H=400 agua/ssDNA. Estos tres estados son compatibles con los resultados obtenidos en experimentos previos, en los que se analizó la variación de la tensión superficial de una monocapa análoga en función de su humedad relativa. Realizando un detallado análisis energético y estructural, hemos podido relacionar las interacciones atomísticas agua-ssDNA con la respuesta estructural y mecánica colectiva de la monocapa.

A niveles bajos de hidratación, la alta afinidad agua-ssDNA favorece el endurecimiento de las hebras de ssDNA y que se aproximen entre ellas, dando lugar a que la deflexión de la monocapa sea más intensa (más grande) a medida que la hidratación aumenta. Cuando 100<H<400 agua/ssDNA (2), la afinidad agua-agua comienza a ser relevante. Esto lleva a un aumento del volumen de los nanoporos, es decir, a que la longitud de las hebras aumente, lo que dificulta el incremento de la curvatura. Cuando H>400 agua/ssDNA (3), los nanoporos están ya completamente llenos de moléculas de agua, de modo que las nuevas moléculas de agua tienen que quedarse en la parte exterior de la monocapa y por lo tanto su incremento no afecta a las propiedades estructurales y mecánicas de la monocapa. Hemos completado este estudio analizando como influye la reducción de la densidad de hebras en la variacion de la respuesta mecánica con la hidratación anteriormente descrita. Los resultados obtenidos confirman que la reducción de la densidad hace que las hebras sigan una nueva tendencia estructural a niveles bajos de hidratación como consecuencia de la existencia de un mayor espaciado entre hebras. Esta nueva tendencia estructural cambia significativamente la respuesta mecánica de la monocapa a niveles de hidratación bajos (H<150 agua/ssDNA). En definitiva, los resultados presentados en este capítulo revelan que las moléculas de agua tienen un rol activo en la estructura y propiedades mecánicas de estas monocapas de ssDNA.

En el cuarto capítulo, hemos explorado el efecto de las modificaciones "átomo a átomo" en la estructura y dinámica de la proteína Azurina. La motivación de este trabajo ha sido no sólo caracterizar el efecto de estas mutaciones en la dinámica de una proteína, sino también entender el cambio en el proceso de transporte electrónico a través de una unión oro-Azurina-oro observado experimentalmente tras reemplazar uno de sus amino-acidos por una cisteína. Simulando la dinámica libre en disolución de tres mutantes de esta proteína análogos a los utilizados en los experimentos, es decir, en las que un aminoacido es remplazado por una cisteína, demostramos que la introducción de mutaciones no afecta prácticamente a la estructura de la Azurina, pero reduce su fluctuación global un $8\,\%$ (se vuelve más rígida). Para analizar como influye este endurecimiento en la interacción Azurina-oro, hemos simulado el proceso de adsorción

sobre una superficie de oro de la forma "wild-type" y mutada de la Azurina. Para obtener una variedad estadística más amplia del proceso de adsorción, hemos estudiado este proceso partiendo de cuatro orientaciones proteina-oro diferentes. Nuestras simulaciones han mostrado que la forma "wild-type" de la proteína tiene una alta mobilidad durante la adsorción, pues experimenta durante este proceso grandes reorientaciones para finalmente adsorberse en una de las dos configuraciones siguientes: en horizontal, con las cisteínas en contacto con la superficie, o a través del parche hidrofóbico. Por el contrario, la forma mutante muestra una menor habilidad de reorientación, dado que siempre se adsorbe con la misma orientación con la que empezó al incio de la simulación. Esta diferencia en el comportamiento de ambas formas sugiere que restringir las fluctuaciones de la Azurina mediante la introducción de mutaciones puede alterar significativamente su proceso de adsorción en oro, modificando incluso su configuración de adsorción. Estos resultados sientan bases para la comprensión de los resultados experimentales, dado que variaciones en la configuración de la Azurina respecto al oro pueden, en última estancia, afectar la conductividad de la Azurina en la uniónes oro-Azurina-oro.

En el quinto capítulo, hemos explorado el efecto de las modificaciones "átomo a átomo" en las propiedades termo-mecánicas del grafeno. Este trabajo combina simulaciones de Dinámica Molecular con resultados experimentales, ya que ha sido realizado en estrecha colaboración con un grupo experimental de nuestra universidad. Nuestro objetivo era revelar la influencia de una densidad pequeña de monovacantes en las fluctuaciones térmicas fuera de plano del grafeno. Estas fluctuationes causan que el grafeno se contraiga cuando se calienta, es decir, el coeficiente de expansión térmica del grafeno (TEC) es negativo. Considerando esta relación entre el TEC y las fluctuaciones fuera de plano, decidimos estudiar la influencia de la presencia de defectos en estas fluctuaciones mediante la estimacion del TEC del grafeno (tanto experimentalmente como con simulaciones) con diferentes densidades de monovacantes (de 0 a 6 $\times 10^{12}$ cm$^{-2}$). Los experimentos han demostrado que el TEC del grafeno se modifica con la introducción de estas monovacantes, reduciéndolo casi a cero cuando la distancia mínima entre vacantes es de $\sim$ 4nm. Por medio de simulaciones de dinámica molecular

a presion constante, hemos podido cuantificar las variaciones en el parámetro de red, con una precisión de $10^{-5}$Å, permitiéndonos estimar con precision el TEC negativo del grafeno, y confirmar que se reduce con la introducción de monovacantes. Para comprender esta reducción, hemos analizado la distribución de tensión a lo largo de las láminas de grafeno, por medio de simulaciones a volumen constante. Esto nos ha permitido concluir que la introducción de monovacantes crea campos de tensión en las proximidades de las mismas, lo que afecta directamente a los modos fuera de plano.

En resumen, este trabajo presenta un estudio amplio del papel de las interacciones atomísticas en las propiedades dinámicas de sistemas en la nanoescala. En particular, hemos revelado los mecanismos atómicos con los que el solvente (el agua) interviene en el proceso de adsorción de una proteína sobre una superficie hidrófoba. Además, hemos desvelado que el agua juega un papel activo en la estructura y respuesta mecánica de sistemas biológicos complejos como los agregados de ssDNAs. Por último, hemos demostrado que la introducción de modificaciones "átomo a átomo" en la estructura intrínseca de dos nano-sistemas diferentes (la proteína Azurina, y el grafeno), cambia significativamente sus propiedades mecánicas, justificando también los mecanismos físicos que subyacen a este cambio.

# List of publications by author

### Publications related to this thesis

- Guillermo López-Polín, Maria Ortega, J. G. Vilhena, Irene Alda, J.Gomez-Herrero, Pedro A. Serena, C. Gomez-Navarro, Rubén Pérez, "Tailoring the thermal expansion of graphene via controlled defect creation", Carbon **116**, 670 (2017).

- Marta P. Ruiz, Albert C. Aragonès, Nuria Camarero, J. G. Vilhena, Maria Ortega,Linda A. Zotti, Rubén Pérez, Juan Carlos Cuevas, Pau Gorostiza, Ismael Díez-Pérez "Bioengineering a Single-Protein Junction", J. Am. Chem. Soc. **139**, 15337 (2017).

- Maria Ortega, J. G. Vilhena, Pamela Rubio-Pereda, P.A. Serena and Rubén Pérez "Assessing the Accuracy of Different Solvation Models To Describe Protein Adsorption", J. Chem. Theory Comput. **15**, 2548 (2019).

- Maria Ortega, J. G. Vilhena, Linda A. Zotti, Ismael Díez-Pérez, Juan Carlos Cuevas and Rubén Pérez "Tuning Structure and Dynamics of Blue Copper Azurin Junctions via Single Amino-Acid Mutations", Biomolecules **9**, 611 (2019).

- Maria Ortega, J. G. Vilhena and Rubén Pérez "Tuning the Mechanical Response of single-stranded DNA (ssDNA) films via its hydration level: insights from MD Simulations", **in preparation.**

**Publications not related to this thesis**

- Carlos Romero-Muñiz, Maria Ortega, J. G. Vilhena, I. Díez-Pérez, Juan Carlos Cuevas, Rubén Pérez and Linda A. Zotti "Ab initio electronic structure calculations of entire blue copper azurins", Phys. Chem. Chem. Phys. **20**, 30392 (2018).

- Carlos Romero-Muñiz, Maria Ortega, J. G. Vilhena, Ismael Diéz-Pérez, Juan Carlos Cuevas, Rubén Pérez and Linda A. Zotti "Mechanical Deformation and Electronic Structure of a Blue Copper Azurin in a Solid-State Junction", Biomolecules **9**, 506 (2019).

# Bibliography


[1] N. R. Council, "The importance of nanoscale science and technology," in *Small wonders, endless frontiers: A review of the National Nanotechnology Initiative* (National Academies Press, 2002) Chap. 1.

[2] M. B. Cortie, Gold Bulletin **37**, 12 (2004).

[3] M. Castellanos, R. Pérez, C. Carrasco, M. Hernando-Pérez, J. Gómez-Herrero, P. J. de Pablo, and M. G. Mateu, Proceedings of the National Academy of Sciences **109**, 12028 (2012).

[4] G. López-Polín, C. Gómez-Navarro, V. Parente, F. Guinea, M. Katsnelson, F. Pérez-Murano, and J. Gómez-Herrero, Nature Physics **11**, 26 (2014).

[5] P. Jungwirth, The Journal of Physical Chemistry Letters **6**, 2449 (2015).

[6] P. Ren, J. Chun, D. G. Thomas, M. J. Schnieders, M. Marucho, J. Zhang, and N. A. Baker, Quarterly Reviews of Biophysics **45**, 427 (2012).

[7] R. Luo, L. David, and M. K. Gilson, Journal of Computational Chemistry **23**, 1244 (2002).

[8] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, Journal of the American Chemical Society **112**, 6127 (1990).

[9] A. Onufriev, D. A. Case, and D. Bashford, Journal of Computational Chemistry **23**, 1297 (2002).

[10] M. Feig, A. Onufriev, M. Lee, W. Im, D. A. Case, and C. L. Brooks, Journal of Computational Chemistry **25**, 265 (2004).

[11] P. Ferrara, J. Apostolakis, and A. Caflisch, Proteins: Structure, Function, and Bioinformatics **46**, 24 (2002).

[12] M. P. Ruiz, A. C. Aragonès, N. Camarero, J. G. Vilhena, M. Ortega, L. A. Zotti, R. Pérez, J. C. Cuevas, P. Gorostiza, and I. Díez-Pérez, Journal of the American Chemical Society **139**, 15337 (2017).

[13] L. Vicarelli, S. J. Heerema, C. Dekker, and H. W. Zandbergen, ACS Nano **9**, 3428 (2015).

[14] H. González-Herrero, J. M. Gómez-Rodríguez, P. Mallet, M. Moaied, J. J. Palacios, C. Salgado, M. M. Ugeda, J.-Y. Veuillen, F. Yndurain, and I. Brihuega, Science **352**, 437 (2016).

[15] P. Fiorani, A. Bruselles, M. Falconi, G. Chillemi, A. Desideri, and P. Benedetti, Journal of Biological Chemistry **278**, 43268 (2003).

[16] G. Rajasekaran, P. Narayanan, and A. Parashar, Critical Reviews in Solid State and Materials Sciences **41**, 47 (2016).

[17] M. Sammalkorpi, A. Krasheninnikov, A. Kuronen, K. Nordlund, and K. Kaski, Physical Review B **70**, 245416 (2004).

[18] Z. Han and A. Fina, Progress in Polymer Science **36**, 914 (2011).

[19] A. Kumar, K. Sharma, P. K. Singh, and V. K. Dwivedi, Materials Today: Proceedings **4**, 4013 (2017).

[20] N. Jing, Q. Xue, C. Ling, M. Shan, T. Zhang, X. Zhou, and Z. Jiao, RSC Advances **2**, 9124 (2012).

[21] N. Chandra, S. Namilae, and C. Shet, Physical Review B **69**, 094101 (2004).

[22] C. Gómez-Navarro, M. Burghard, and K. Kern, Nano Letters **8**, 2045 (2008).

[23] C. S. Ruiz-Vargas, H. L. Zhuang, P. Y. Huang, A. M. van der Zande, S. Garg, P. L. McEuen, D. A. Muller, R. G. Hennig, and J. Park, Nano Letters **11**, 2259 (2011).

[24] R. Ansari, S. Ajori, and B. Motevalli, Superlattices and Microstructures **51**, 274 (2012).

[25] B. Rizzuti, L. Sportelli, and R. Guzzi, Biophysical Chemistry **94**, 107 (2001).

[26] C. G. P. Doss, B. Rajith, N. Garwasis, P. R. Mathew, A. S. Raju, K. Apoorva, D. William, N. Sadhana, T. Himani, and I. Dike, Applied and Translational Genomics **1**, 37 (2012).

[27] K. Gekko, Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology **1595**, 382 (2002).

[28] A. Kumar, V. Rajendran, R. Sethumadhavan, and R. Purohit, Journal of Biomolecular Structure and Dynamics **32**, 394 (2014).

[29] Y. Fan, A. Cembran, S. Ma, and J. Gao, Biochemistry **52**, 2036 (2013).

[30] E. Behmard, P. Abdolmaleki, E. B. Asadabadi, and S. Jahandideh, Journal of Biomolecular Structure and Dynamics **29**, 379 (2011).

[31] P. J. P. Carrillo, M. Medrano, A. Valbuena, A. Rodríguez-Huete, M. Castellanos, R. Pérez, and M. G. Mateu, ACS Nano **11**, 2194 (2017).

[32] W. F. van Gunsteren and H. J. C. Berendsen, Angewandte Chemie International Edition **29**, 992 (1990).

[33] A. Hospital, J. R. Goñi, M. Orozco, and J. L. Gelpí, Advances and applications in bioinformatics and chemistry: AABC **8**, 37 (2015).

[34] M. González, École thématique de la Société Française de la Neutronique **12**, 169 (2011).

[35] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, Journal of Computational Chemistry **26**, 1668 (2005).

[36] R. Salomon-Ferrer, D. A. Case, and R. C. Walker, Wiley Interdisciplinary Reviews: Computational Molecular Science **3**, 198 (2013).

[37] H. Lei, C. Wu, H. Liu, and Y. Duan, Proceedings of the National Academy of Sciences **104**, 4925 (2007).

[38] Q. Shao and W. Zhu, Physical Chemistry Chemical Physics **21**, 11924 (2019).

[39] A. Marin-Gonzalez, J. G. Vilhena, R. Perez, and F. Moreno-Herrero, Proceedings of the National Academy of Sciences **114**, 7049 (2017).

[40] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, Journal of Chemical Theory and Computation **11**, 3696 (2015).

[41] S. J. Stuart, A. B. Tutein, and J. A. Harrison, The Journal of Chemical Physics **112**, 6472 (2000).

[42] D. W. Brenner, Physical Review B **42**, 9458 (1990).

[43] J. Tersoff, Physical Review B **37**, 6991 (1988).

[44] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, The Journal of Physical Chemistry A **105**, 9396 (2001).

[45] D. Kvashnin and P. Sorokin, The Journal of Physical Chemistry Letters **6**, 2384 (2015).

[46] W. Gao and R. Huang, Journal of the Mechanics and Physics of Solids **66**, 42 (2014).

[47] N. C. B. Mostério and A. F. Fonseca, Physical Review B **89**, 195437 (2014).

[48] R. J. Loncharich, B. R. Brooks, and R. W. Pastor, Biopolymers **32**, 523 (1992).

[49] G. J. Martyna, D. J. Tobias, and M. L. Klein, The Journal of Chemical Physics **101**, 4177 (1994).

[50] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, The Journal of Chemical Physics **81**, 3684 (1984).

[51] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, 2nd ed. (Oxford University Press, 2017).

[52] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, and K. Schulten, "Steered molecular dynamics," in *Computational Molecular Dynamics: Challenges, Methods, Ideas* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999) pp. 39–65.

[53] M. Feig and C. L. Brooks, Current Opinion in Structural Biology **14**, 217 (2004).

[54] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, The Journal of Physical Chemistry **100**, 19824 (1996).

[55] A. Onufriev, D. Bashford, and D. A. Case, Proteins **55**, 383 (2004).

[56] J. G. Vilhena, A. C. Dumitru, E. T. Herruzo, J. I. Mendieta-Moreno, R. Garcia, P. A. Serena, and R. Perez, Nanoscale **8**, 13463 (2016).

[57] W. Norde, T. A. Horbett, and J. L. Brash, "Proteins at interfaces iii: Introductory overview," in *Proteins at Interfaces III State of the Art*, ACS Symposium Series, Vol. 1120 (American Chemical Society, 2012) Chap. 1, pp. 1–34.

[58] M. Rabe, D. Verdes, and S. Seeger, Advances in Colloid and Interface Science **162**, 87 (2011).

[59] W. R. Sanhai, J. H. Sakamoto, R. Canady, and M. Ferrari, Nature Nanotechnology **3**, 242 (2008).

[60] Z. Zhang, M. Zhang, S. Chen, T. A. Horbetta, B. D. Ratner, and S. Jiang, Biomaterials **29**, 4285 (2008).

[61] A. A. Yanik, M. Huang, O. Kamohara, A. Artar, T. W. Geisbert, J. H. Connor, and H. Altug, Nano Letters **10**, 4962 (2010).

[62] E. T. Herruzo, H. Asakawa, T. Fukuma, and R. Garcia, Nanoscale **5**, 2678 (2013).

[63] K. Kuchuk and U. Sivan, Nano Letters **18**, 2733 (2018).

[64] F. Höök, J. Vörös, M. Rodahl, R. Kurrat, P. Böni, J. Ramsden, M. Textor, N. Spencer, P. Tengvall, J. Gold, and B. Kasemo, Colloids and Surfaces B: Biointerfaces **24**, 155 (2002).

[65] M. Geisler, S. Xiao, E. M. Puchner, F. Gräter, and T. Hugel, Journal of the American Chemical Society **132**, 17277 (2010).

[66] B. Pelaz, P. del Pino, P. Maffre, R. Hartmann, M. Gallego, S. Rivera-Fernández, J. M. de la Fuente, G. U. Nienhaus, and W. J. Parak, ACS Nano **9**, 6996 (2015).

[67] I. Sharma and S. K. Pattanayek, Biophysical Chemistry **226**, 14 (2017).

[68] R. A. Latour, Biointerphases **3**, FC2 (2008).

[69] M. J. Penna, M. Mijajlovic, and M. J. Biggs, Journal of the American Chemical Society **136**, 5323 (2014).

[70] N. T. Southall, K. A. Dill, and A. D. J. Haymet, The Journal of Physical Chemistry B **106**, 521 (2002).

[71] A. Biela, N. N. Nasief, M. Betz, A. Heine, D. Hangauer, and G. Klebe, Angewandte Chemie International Edition **52**, 1822 (2013).

[72] J. J. Gray, Current Opinion in Structural Biology **14**, 110 (2004).

[73] R. A. Latour, "Molecular simulation of protein-surface interactions," in *Biological Interactions on Materials Surfaces: Understanding and Controlling Protein, Cell, and Tissue Responses* (Springer US, New York, NY, 2009) pp. 69–95.

[74] M. Orozco and F. J. Luque, Chemical Reviews **100**, 4187 (2000).

[75] H. Gohlke and D. A. Case, Journal of Computational Chemistry **25**, 238 (2004).

[76] B. D. Bursulaya and C. L. Brooks, The Journal of Physical Chemistry B **104**, 12378 (2000).

[77] J. Kleinjung and F. Fraternali, Current Opinion in Structural Biology **25**, 126 (2014).

[78] B. N. Dominy and C. L. Brooks, Journal of Computational Chemistry **23**, 147 (2002).

[79] Y. A. Arnautova, Y. N. Vorobjev, J. A. Vila, and H. A. Scheraga, Proteins: Structure, Function, and Bioinformatics **77**, 38 (2009).

[80] C. J. Cramer and D. G. Truhlar, Chemical Reviews **99**, 2161 (1999).

[81] R. Garcia and E. T. Herruzo, Nature Nanotechnology **7**, 217 (2012).

[82] Z. Yu, M. P. Jacobson, J. Josovitz, C. S. Rapp, and R. A. Friesner, The Journal of Physical Chemistry B **108**, 6643 (2004).

[83] Q. Cui, The Journal of Chemical Physics **117**, 4720 (2002).

[84] M. S. Lee, F. R. Salsbury, and M. A. Olson, Journal of Computational Chemistry **25**, 1967 (2004).

[85] O. M. Szklarczyk, N. S. Bieler, P. H. Hünenberger, and W. F. van Gunsteren, Journal of Chemical theory and Computation **11**, 5447 (2015).

[86] G. Raffaini and F. Ganazzoli, Langmuir **29**, 4883 (2013).

[87] G. Raffaini and F. Ganazzoli, Langmuir **26**, 5679 (2010).

[88] G. Raffaini and F. Ganazzoli, Macromolecular Bioscience **7**, 552 (2007).

[89] F. Ganazzoli and G. Raffaini, Physical Chemistry Chemical Physics **7**, 3651 (2005).

[90] G. Raffaini and F. Ganazzoli, Langmuir **19**, 3403 (2003).

[91] C. Mücksch and H. M. Urbassek, Chemical Physics Letters **510**, 252 (2011).

[92] C. Mücksch and H. M. Urbassek, PLoS ONE **8**, e64883 (2013).

[93] D. Horinek, A. Serr, M. Geisler, T. Pirzer, U. Slotta, S. Q. Lud, J. A. Garrido, T. Scheibel, T. Hugel, and R. R. Netz, Proceedings of the National Academy of Sciences **105**, 2842 (2008).

[94] T. Wei, M. A. Carignano, and I. Szleifer, Langmuir **27**, 12074 (2011).

[95] L. Ou, Y. Luo, and G. Wei, The Journal of Physical Chemistry B **115**, 9813 (2011).

[96] T. Wei, M. A. Carignano, and I. Szleifer, The Journal of Physical Chemistry B **116**, 10189 (2012).

[97] J. G. Vilhena, P. Rubio-Pereda, P. Vellosillo, P. A. Serena, and R. Pérez, Langmuir **32**, 1742 (2016).

[98] K. Grattendick and S. Pross, "Immunoglobulins," in *xPharm: The Comprehensive Pharmacology Reference* (Elsevier, New York, 2007) pp. 1 – 6.

[99] L. Beaudet, R. Rodriguez-Suarez, M.-H. Venne, M. Caron, J. Bédard, V. Brechler, S. Parent, and M. Bielefeld-Sévigny, Nature Methods **5**, an8 (2008).

[100] J. Tamayo, P. M. Kosaka, J. J. Ruz, San Paulo, and M. Calleja, Chemical Society Reviews **42**, 1287 (2013).

[101] S. Patil, N. F. Martinez, J. R. Lozano, and R. Garcia, Journal of Molecular Recognition **20**, 516 (2007).

[102] J. Preiner, N. S. Losilla, A. Ebner, P. Annibale, F. Biscarini, R. Garcia, and P. Hinterdorfer, Nano Letters **9**, 571 (2009).

[103] H. Xu, J. R. Lu, and D. E. Williams, The Journal of Physical Chemistry B **110**, 1907 (2006).

[104] L. J. Harris, S. B. Larson, K. W. Hasel, and A. McPherson, Biochemistry **36**, 1581 (1997).

[105] J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, and A. Onufriev, Nucleic Acids Research **33**, W368 (2005).

[106] M. L. Connolly, Journal of Applied Crystallography **16**, 548 (1983).

[107] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins: Structure, Function, and Bioinformatics **78**, 1950 (2010).

[108] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, and R. J. Woods, Journal of Computational Chemistry **29**, 622 (2008).

[109] J. P. Brandt, T. W. Patapoff, and S. R. Aragon, Biophysical Journal **99**, 905 (2010).

[110] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, Journal of Computational Chemistry **25**, 1157 (2004).

[111] J.-L. Tsai and J.-F. Tu, Materials and Design **31**, 194 (2010).

[112] M. Pykal, P. Jurecka, F. Karlicky, and M. Otyepka, Physical Chemistry Chemical Physics **18**, 6351 (2016).

[113] J. G. Vilhena, C. Pimentel, P. Pedraz, F. Luo, P. A. Serena, C. M. Pina, E. Gnecco, and R. Pérez, ACS Nano **10**, 4288 (2016).

[114] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, The Journal of Chemical Physics **79**, 926 (1983).

[115] A. Onufriev, D. Bashford, and D. A. Case, The Journal of Physical Chemistry B **104**, 3712 (2000).

[116] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, Chemical Physics Letters **246**, 122 (1995).

[117] H. Liu, F. Chen, H. Sun, D. Li, and T. Hou, Journal of Chemical Theory and Computation **13**, 1827 (2017).

[118] A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, and R. C. Walker, Journal of Chemical Theory and Computation **8**, 1542 (2012).

[119] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker, Journal of Chemical Theory and Computation **9**, 3878 (2013).

[120] M. Crowley, T. Darden, T. Cheatham, and D. Deerfield, The Journal of Supercomputing **11**, 255 (1997).

[121] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, The Journal of Chemical Physics **103**, 8577 (1995).

[122] S. Miyamoto and P. A. Kollman, Journal of Computational Chemistry **13**, 952 (1992).

[123] S. Rogge, L. Vanduyfhuys, A. Ghysels, M. Waroquier, T. Verstraelen, G. Maurin, and V. Van Speybroeck, Journal of Chemical Theory and Computation **11**, 5583 (2015).

[124] D. R. Roe and T. E. Cheatham, Journal of Chemical Theory and Computation **9**, 3084 (2013).

[125] W. Kabsch and C. Sander, Biopolymers **22**, 2577 (1983).

[126] R. P. Joosten, T. A. H. te Beek, E. Krieger, M. L. Hekkelman, R. W. W. Hooft, R. Schneider, C. Sander, and G. Vriend, Nucleic Acids Research **39**, D411 (2010).

[127] A. W. Vermeer, W. Norde, and A. van Amerongen, Biophysical Journal **79**, 2150 (2000).

[128] A. P. A. Raju, S. C. Offerman, P. Gorgojo, C. Valles, E. V. Bichenkova, H. S. Aojula, A. Vijayraghavan, R. J. Young, K. S. Novoselov, I. A. Kinloch, and D. J. Clarke, RSC Advances **6**, 69551 (2016).

[129] R. Garrett and C. Grisham, *Biochemistry* (Cengage Learning, 2012).

[130] J. Mertens, C. Rogero, M. Calleja, D. Ramos, J. A. Martín-Gago, C. Briones, and J. Tamayo, Nature Nanotechnology **3**, 301 (2008).

[131] D. Martin-Jimenez, E. Chacon, P. Tarazona, and R. Garcia, Nature Communications **7**, 12164 (2016).

[132] D. Martin-Jimenez and R. Garcia, The Journal of Physical Chemistry Letters **8**, 5707 (2017).

[133] M. Ricci, P. Spijker, and K. Voïtchovsky, Nature Communications **5**, 4400 (2014).

[134] K. Voïtchovsky and M. Ricci, in *Colloidal Nanocrystals for Biomedical Applications VII*, Vol. 8232 (2012) pp. 120 – 127.

[135] D. Ortiz-Young, H.-C. Chiu, S. Kim, K. Voïtchovsky, and E. Riedo, Nature Communications **4**, 2482 (2013).

[136] K. Falk, F. Sedlmeier, L. Joly, R. R. Netz, and L. Bocquet, Nano Letters **10**, 4067 (2010).

[137] C. Sendner, D. Horinek, L. Bocquet, and R. R. Netz, Langmuir **25**, 10768 (2009).

[138] D. M. Huang, C. Sendner, D. Horinek, R. R. Netz, and L. Bocquet, Physical Review Letters **101**, 226101 (2008).

[139] M. Ortega, J. G. Vilhena, P. Rubio-Pereda, P. A. Serena, and R. Pérez, Journal of Chemical Theory and Computation **15**, 2548 (2019).

[140] P. C. Nalam, J. N. Clasohm, A. Mashaghi, and N. D. Spencer, Tribology Letters **37**, 541 (2010).

[141] P. C. Nalam, S. N. Ramakrishna, R. M. Espinosa-Marzal, and N. D. Spencer, Langmuir **29**, 10149 (2013).

[142] M. Steenackers, A. Küller, N. Ballav, M. Zharnikov, M. Grunze, and R. Jordan, Small **3**, 1764 (2007).

[143] A. Nomura, K. Ohno, T. Fukuda, T. Sato, and Y. Tsujii, Polymer Chemistry **3**, 148 (2012).

[144] S. N. Ramakrishna, R. M. Espinosa-Marzal, V. V. Naik, P. C. Nalam, and N. D. Spencer, Langmuir **29**, 15251 (2013).

[145] M. K. Singh, C. Kang, P. Ilg, R. Crockett, M. Kröger, and N. D. Spencer, Macromolecules **51**, 10174 (2018).

[146] S. Desseaux, J. P. Hinestrosa, N. Schüwer, B. S. Lokitz, J. F. Ankner, S. M. Kilbey, K. Voitchovsky, and H.-A. Klok, Macromolecules **49**, 4609 (2016).

[147] X. Liu, Q. Ye, B. Yu, Y. Liang, W. Liu, and F. Zhou, Langmuir **26**, 12377 (2010).

[148] M. K. Singh, P. Ilg, R. M. Espinosa-Marzal, M. Kröger, and N. D. Spencer, Langmuir **31**, 4798 (2015).

[149] M. A. C. Stuart, W. T. S. Huck, J. Genzer, M. Müller, C. Ober, M. Stamm, G. B. Sukhorukov, I. Szleifer, V. V. Tsukruk, M. Urban, F. Winnik, S. Zauscher, I. Luzinov, and S. Minko, Nature Materials **9**, 101 (2010).

[150] J. Klein, Y. Kamiyama, H. Yoshizawa, J. N. Israelachvili, G. H. Fredrickson, P. Pincus, and L. J. Fetters, Macromolecules **26**, 5552 (1993).

[151] C. H. Mathis, R. Simič, C. Kang, S. N. Ramakrishna, L. Isa, and N. D. Spencer, Polymer **169**, 115 (2019).

[152] G. L. Kenausis, J. Vörös, D. L. Elbert, N. Huang, R. Hofer, L. Ruiz-Taylor, M. Textor, J. A. Hubbell, and N. D. Spencer, The Journal of Physical Chemistry B **104**, 3298 (2000).

[153] M. P. Weir and A. J. Parnell, Polymers **3**, 2107 (2011).

[154] R. M. Espinosa-Marzal, P. C. Nalam, S. Bolisetty, and N. D. Spencer, Soft Matter **9**, 4045 (2013).

[155] A. Ghosh and M. Bansal, Acta Crystallographica Section D **59**, 620 (2003).

[156] P. S. Gil, D. J. Lacks, P. Parisse, L. Casalis, and M. D. Nkoua Ngavouka, Soft Matter **14**, 9675 (2018).

[157] J. Fritz, M. K. Baller, H. P. Lang, H. Rothuizen, P. Vettiger, E. Meyer, H. J. Güntherodt, C. Gerber, and J. K. Gimzewski, Science **288**, 316 (2000).

[158] R. McKendry, J. Zhang, Y. Arntz, T. Strunz, M. Hegner, H. P. Lang, M. K. Baller, U. Certa, E. Meyer, H.-J. Güntherodt, and C. Gerber, Proceedings of the National Academy of Sciences **99**, 9783 (2002).

[159] N. Zhang, Z. Tan, J. Li, W. Meng, and L. Xu, Current Opinion in Colloid and Interface Science **16**, 592 (2011).

[160] C. M. Dominguez, P. M. Kosaka, G. Mokry, V. Pini, O. Malvar, M. del Rey, D. Ramos, San Paulo, J. Tamayo, and M. Calleja, Langmuir **30**, 10962 (2014).

[161] C. M. Domínguez, D. Ramos, J. I. Mendieta-Moreno, J. L. G. Fierro, J. Mendieta, J. Tamayo, and M. Calleja, Scientific Reports **7**, 536 (2017).

[162] M. Wagman, S. Medalion, and Y. Rabin, Macromolecules **45**, 9517 (2012).

[163] Z. Xin-Jun, G. Zhi-Fu, and J. Zhong-Ying, Chinese Physics B **24**, 044701 (2015).

[164] S. Choudhuri, "Fundamentals of genes and genomes," in *Bioinformatics for Beginners* (Academic Press, Oxford, 2014) Chap. 1, pp. 1 – 25.

[165] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. C. III, C. A. Laughton, and M. Orozco, Biophysical Journal **92**, 3817 (2007).

[166] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, Journal of the American Chemical Society **118**, 2309 (1996).

[167] A. T. Guy, T. J. Piggot, and S. Khalid, Biophysical Journal **103**, 1028 (2012).

[168] M. Nakano, H. Tateishi-Karimata, S. Tanaka, F. Tama, O. Miyashita, S. ichi Nakano, and N. Sugimoto, Chemical Physics Letters **660**, 250 (2016).

[169] J. G. Vilhena, E. Gnecco, R. Pawlak, F. Moreno-Herrero, E. Meyer, and R. Pérez, The Journal of Physical Chemistry B **122**, 840 (2018).

[170] J. E. Andrews, S. Sinha, P. W. Chung, and S. Das, Physical Chemistry Chemical Physics **18**, 23482 (2016).

[171] M. Kabeláč, O. Kroutil, M. Předota, F. Lankaš, and M. Šíp, Physical Chemistry Chemical Physics **14**, 4217 (2012).

[172] H. S. Kim, S. M. Huang, and Y. G. Yingling, MRS Advances **1**, 1883 (2016).

[173] S. K. Min, W. Y. Kim, Y. Cho, and K. S. Kim, Nature Nanotechnology **6**, 162 (2011).

[174] I. S. Joung and T. E. Cheatham, The Journal of Physical Chemistry B **113**, 13279 (2009).

[175] G. G. Stoney, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **82**, 172 (1909).

[176] X. Feng, y.-s. Huang, and A. J. Rosakis, Journal of Applied Mechanics **74**, 1276 (2007).

[177] N. J. Tao, S. M. Lindsay, and A. Rupprecht, Biopolymers **28**, 1019 (1989).

[178] N. Lavalle, S. A. Lee, and A. Rupprecht, Biopolymers **30**, 877 (1990).

[179] J. Camunas-Soler, M. Ribezzi-Crivellari, and F. Ritort, Annual Review of Biophysics **45**, 65 (2016).

[180] K. Rechendorff, G. Witz, J. Adamcik, and G. Dietler, The Journal of Chemical Physics **131**, 095103 (2009).

[181] D. V. Klinov, T. V. Neretina, V. V. Prokhorov, T. V. Dobrynina, K. G. Aldarov, and V. V. Demin, Biochemistry (Moscow) **74**, 1150 (2009).

[182] E. Westhof, Annual Review of Biophysics and Biophysical Chemistry **17**, 125 (1988).

[183] E. Westhof, "Structural water bridges in nucleic acids," in *Water and Biological Macromolecules* (Macmillan Education UK, London, 1993) pp. 226–243.

[184] S. Cui, C. Albrecht, F. Kühner, and H. E. Gaub, Journal of the American Chemical Society **128**, 6636 (2006).

[185] F. Oesterhelt, M. Rief, and H. E. Gaub, New Journal of Physics **1**, 6 (1999).

[186] G. Wu, H. Ji, K. Hansen, T. Thundat, R. Datar, R. Cote, M. F. Hagan, A. K. Chakraborty, and A. Majumdar, Proceedings of the National Academy of Sciences **98**, 1560 (2001).

[187] J. A. Fereiro, X. Yu, I. Pecht, M. Sheves, J. C. Cuevas, and D. Cahen, Proceedings of the National Academy of Sciences **115**, E4577 (2018).

[188] J. M. Artés, M. López-Martínez, I. Díez-Pérez, F. Sanz, and P. Gorostiza, Small **10**, 2537 (2014).

[189] A. S. Venkat, S. Corni, and R. Di Felice, Small **3**, 1431 (2007).

[190] C. D. Bostick, S. Mukhopadhyay, I. Pecht, M. Sheves, D. Cahen, and D. Lederman, Reports on Progress in Physics **81**, 26601 (2018).

[191] J. E. Huheey, E. A. Keiter, R. L. Keiter, and O. K. Medhi, *Inorganic chemistry: principles of structure and reactivity* (Pearson Education India, 2006).

[192] E. N. Baker, Journal of Molecular Biology **203**, 1071 (1988).

[193] H. Nar, A. Messerschmidt, R. Huber, M. van de Kamp, and G. W. Canters, Journal of Molecular Biology **221**, 765 (1991).

[194] A. Alessandrini, M. Gerunda, G. Canters, M. Verbeet, and P. Facci, Chemical Physics Letters **376**, 625 (2003).

[195] B. Rizzuti, L. Sportelli, and R. Guzzi, Biophysical Chemistry **125**, 532 (2007).

[196] A. R. Bizzarri, C. Baldacchini, and S. Cannistraro, Langmuir **33**, 9190 (2017).

[197] C. Romero-Muñiz, M. Ortega, J. G. Vilhena, I. Diéz-Pérez, J. C. Cuevas, R. Pérez, and L. A. Zotti, Biomolecules **9**, 506 (2019).

[198] C. Baldacchini, V. Kumar, A. R. Bizzarri, and S. Cannistraro, Applied Physics Letters **106**, 183701 (2015).

[199] X. Yu, R. Lovrincic, L. Sepunaru, W. Li, A. Vilan, I. Pecht, M. Sheves, and D. Cahen, ACS Nano **9**, 9955 (2015).

[200] S. Kradolfer, E. Lipiec, C. Baldacchini, A. R. Bizzarri, S. Cannistraro, and R. Zenobi, ACS Nano **11**, 12824 (2017).

[201] S. S. Panda, H. E. Katz, and J. D. Tovar, Chemical Society Reviews **47**, 3640 (2018).

[202] Q. Chi, O. Farver, and J. Ulstrup, Proceedings of the National Academy of Sciences **102**, 16203 (2005).

[203] Q. Chi, J. Zhang, J. E. T. Andersen, and J. Ulstrup, The Journal of Physical Chemistry B **105**, 4669 (2001).

[204] W. Haiss, R. J. Nichols, H. van Zalinge, S. J. Higgins, D. Bethell, and D. J. Schiffrin, Physical Chemistry Chemical Physics **6**, 4330 (2004).

[205] W. Haiss, C. Wang, I. Grace, A. S. Batsanov, D. J. Schiffrin, S. J. Higgins, M. R. Bryce, C. J. Lambert, and R. J. Nichols, Nature Materials **5**, 995 (2006).

[206] I. Díez-Pérez, J. Hihath, Y. Lee, L. Yu, L. Adamska, M. A. Kozhushner, I. I. Oleynik, and N. Tao, Nature Chemistry **1**, 635 (2009).

[207] A. C. Aragonès, N. Darwish, S. Ciampi, F. Sanz, J. J. Gooding, and I. Díez-Pérez, Nature Communications **8**, 15056 (2017).

[208] E. Leary, M. T. González, C. van der Pol, M. R. Bryce, S. Filippone, N. Martín, G. Rubio-Bollinger, and N. Agraït, Nano Letters **11**, 2236 (2011).

[209] A. C. Aragonès, N. Darwish, W. J. Saletra, L. Pérez-García, F. Sanz, J. Puigmartí-Luis, D. B. Amabilino, and I. Díez-Pérez, Nano Letters **14**, 4751 (2014).

[210] U. Kolczak, C. Dennison, A. Messerschmidt, and G. Canters, "Azurin and azurin mutants," in *Handbook of Metalloproteins* (American Cancer Society, 2006).

[211] D. K. Garner, M. D. Vaughan, H. J. Hwang, M. G. Savelieff, S. M. Berry, J. F. Honek, and Y. Lu, Journal of the American Chemical Society **128**, 15608 (2006).

[212] N. M. Marshall, D. K. Garner, T. D. Wilson, Y.-G. Gao, H. Robinson, M. J. Nilges, and Y. Lu, Nature **462**, 113 (2009).

[213] R. Di Felice, A. Selloni, and E. Molinari, The Journal of Physical Chemistry B **107**, 1151 (2003).

[214] J. S. Kretchmer, N. Boekelheide, J. J. Warren, J. R. Winkler, H. B. Gray, and T. F. Miller, Proceedings of the National Academy of Sciences **115**, 6129 (2018).

[215] C. Romero-Muñiz, M. Ortega, J. G. Vilhena, I. Díez-Pérez, J. C. Cuevas, R. Pérez, and L. A. Zotti, Physical Chemistry Chemical Physics **20**, 30392 (2018).

[216] M. Ozboyaci, D. B. Kokh, S. Corni, and R. C. Wade, Quarterly Reviews of Biophysics **49**, e4 (2016).

[217] M. C. Childers and V. Daggett, Molecular Systems Design and Engineering **2**, 9 (2017).

[218] R. Di Felice and S. Corni, The Journal of Physical Chemistry Letters **2**, 1510 (2011).

[219] C. Arcangeli, A. R. Bizzarri, and S. Cannistraro, Biophysical Chemistry **78**, 247 (1999).

[220] A. R. Bizzarri, Biophysical Chemistry **122**, 206 (2006).

[221] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, Nucleic Acids Research **28**, 235 (2000).

[222] S.-U. Kim, Y. J. Kim, S.-G. Choi, C.-H. Yea, R. P. Singh, J. Min, B.-K. Oh, and J.-W. Choi, Ultramicroscopy **108**, 1390 (2008).

[223] M. van den Bosch, M. Swart, J. G. Snijders, H. J. C. Berendsen, A. E. Mark, C. Oostenbrink, W. F. van Gunsteren, and G. W. Canters, ChemBioChem **6**, 738 (2005).

[224] L. Paltrinieri, M. Borsari, A. Ranieri, G. Battistuzzi, S. Corni, and C. A. Bortolotti, The Journal of Physical Chemistry Letters **4**, 710 (2013).

[225] A. E. M. Beedle, A. Lezamiz, G. Stirnemann, and S. Garcia-Manyes, Nature Communications **6**, 7894 (2015).

[226] L. Zanetti-Polzi, C. A. Bortolotti, I. Daidone, M. Aschi, A. Amadei, and S. Corni, Organic and Biomolecular Chemistry **13**, 11003 (2015).

[227] L. Zanetti-Polzi, S. Corni, I. Daidone, and A. Amadei, Physical Chemistry Chemical Physics **18**, 18450 (2016).

[228] H. Heinz, T.-J. Lin, R. Kishore Mishra, and F. S. Emami, Langmuir **29**, 1754 (2013).

[229] H. Heinz and H. Ramezani-Dakhel, Chemical Society Reviews **45**, 412 (2016).

[230] R. Pawlak, J. G. Vilhena, A. Hinaut, T. Meier, T. Glatzel, A. Baratoff, E. Gnecco, R. Pérez, and E. Meyer, Nature Communications **10**, 685 (2019).

[231] P. Li, B. P. Roberts, D. K. Chakravorty, and K. M. Merz, Journal of Chemical Theory and Computation **9**, 2733 (2013).

[232] S. L. Grand, A. W. Götz, and R. C. Walker, Computer Physics Communications **184**, 374 (2013).

[233] A. Alessandrini, S. Corni, and P. Facci, Physical Chemistry Chemical Physics **8**, 4383 (2006).

[234] J. M. Artés, I. Díez-Pérez, and P. Gorostiza, Nano Letters **12**, 2679 (2012).

[235] V. Rajapandian, V. Hakkim, and V. Subramanian, The Journal of Physical Chemistry B **114**, 8474 (2010).

[236] P. Rubio-Pereda, J. G. Vilhena, N. Takeuchi, P. A. Serena, and R. Pérez, The Journal of Chemical Physics **146**, 214704 (2017).

[237] C. Arcangeli, A. R. Bizzarri, and S. Cannistraro, Biophysical Chemistry **90**, 45 (2001).

[238] J. W. Curtis, Proteins: Structure, Function, and Bioinformatics **35**, 307 (1999).

[239] W. A. Lim and R. T. Sauer, Journal of Molecular Biology **219**, 359 (1991).

[240] G. A. Petsko and D. Ringe, Annual Review of Biophysics and Bioengineering **13**, 331 (1984).

[241] A. R. Bizzarri, E. Brunori, B. Bonanni, and S. Cannistraro, Journal of Molecular Recognition **20**, 122 (2007).

[242] B. Rizzuti, M. Swart, L. Sportelli, and R. Guzzi, Journal of Molecular Modeling **10**, 25 (2004).

[243] H. Nar, A. Messerschmidt, R. Huber, M. van de Kamp, and G. W. Canters, FEBS Letters **306**, 119 (1992).

[244] B. Bonanni, D. Alliata, L. Andolfi, A. Bizzarri, and S. Cannistraro, "Redox metalloproteins on metal surfaces as hybrid systems for bionanodevices: An extensive characterization at the single molecule level," in *Surface science research developments* (Nova Science Publishers, 2005) pp. 1–73.

[245] J. J. Davis and H. A. O. Hill, Chemical Communications **5**, 393 (2002).

[246] A. K. Geim and K. S. Novoselov, Nature Materials **6**, 183 (2007).

[247] M. I. Katsnelson and A. Fasolino, Accounts of chemical research **46**, 97 (2013).

[248] W. Bao, F. Miao, Z. Chen, H. Zhang, W. Jang, C. Dames, and C. N. Lau, Nature Nanotechnology **4**, 562 (2009).

[249] V. Singh, S. Sengupta, H. S. Solanki, R. Dhall, A. Allain, S. Dhara, P. Pant, and M. M. Deshmukh, Nanotechnology **21**, 165204 (2010).

[250] D. Yoon, Y.-W. Son, and H. Cheong, Nano Letters **11**, 3227 (2011).

[251] F. L. Braghin and N. Hasselmann, Physical Review B **82**, 035407 (2010).

[252] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).

[253] A. A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. N. Lau, Nano Letters **8**, 902 (2008).

[254] C. Lee, X. Wei, J. W. Kysar, and J. Hone, Science **321**, 385 (2008).

[255] K. S. Novoselov, V. I. Fal'ko, L. Colombo, P. R. Gellert, M. G. Schwab, and K. Kim, Nature **490**, 192 (2012).

[256] N. Mounet and N. Marzari, Physical Review B **71**, 205214 (2005).

[257] K. V. Zakharchenko, M. I. Katsnelson, and A. Fasolino, Physical Review Letters **102**, 046808 (2009).

[258] K. V. Zakharchenko, J. H. Los, M. I. Katsnelson, and A. Fasolino, Physical Review B **81**, 235439 (2010).

[259] P. L. de Andres, F. Guinea, and M. I. Katsnelson, Physical Review B **86**, 144103 (2012).

[260] J.-W. Jiang, B.-S. Wang, J.-S. Wang, and H. S. Park, Journal of Physics: Condensed Matter **27**, 083001 (2015).

[261] Q. Lu, M. Arroyo, and R. Huang, Journal of Physics D: Applied Physics **42**, 102002 (2009).

[262] J. C. Meyer, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. J. Booth, and S. Roth, Nature **446**, 60 (2007).

[263] A. Laitinen, M. Oksanen, A. Fay, D. Cox, M. Tomi, P. Virtanen, and P. J. Hakonen, Nano Letters **14**, 3009 (2014).

[264] K. I. Bolotin, K. J. Sikes, J. Hone, H. L. Stormer, and P. Kim, Physical Review Letters **101**, 096802 (2008).

[265] J. H. Seol, I. Jo, A. L. Moore, L. Lindsay, Z. H. Aitken, M. T. Pettes, X. Li, Z. Yao, R. Huang, D. Broido, N. Mingo, R. S. Ruoff, and L. Shi, Science **328**, 213 (2010).

[266] S. Wang, M. Tambraparni, J. Qiu, J. Tipton, and D. Dean, Macromolecules **42**, 5251 (2009).

[267] P. R. Shaina, L. George, V. Yadav, and M. Jaiswal, Journal of Physics: Condensed Matter **28**, 085301 (2016).

[268] A. L. C. da Silva, L. Cândido, J. N. T. Rabelo, G.-Q. Hai, and F. M. Peeters, EPL (Europhysics Letters) **107**, 56004 (2014).

[269] H. H. P. Garza, E. W. Kievit, G. F. Schneider, and U. Staufer, Nano Letters **14**, 4107 (2014).

[270] X.-H. Zha, R.-Q. Zhang, and Z. Lin, The Journal of Chemical Physics **141**, 064705 (2014).

[271] M. Z. Islam, M. Mahboob, L. R. Lowe, and E. S. Bechtel, Journal of Physics D: Applied Physics **46**, 435302 (2013).

[272] D. Nelson, T. Piran, and S. Weinberg, *Statistical Mechanics of Membranes and Surfaces* (World Scientific Publishing, 2004).

[273] J. P. Mathew, R. N. Patel, A. Borah, R. Vijay, and M. M. Deshmukh, Nature Nanotechnology **11**, 747 (2016).

[274] G. Tocci, L. Joly, and A. Michaelides, Nano Letters **14**, 6872 (2014).

[275] I. Horcas, R. Fernández, J. M. Gómez-Rodríguez, J. Colchero, J. Gómez-Herrero, and A. M. Baro, Review of Scientific Instruments **78**, 013705 (2007).

[276] M. R. Begley and T. J. Mackin, Journal of the Mechanics and Physics of Solids **52**, 2005 (2004).

[277] D. Garcia-Sanchez, A. M. van der Zande, A. S. Paulo, B. Lassagne, P. L. McEuen, and A. Bachtold, Nano Letters **8**, 1399 (2008).

[278] O. Lehtinen, J. Kotakoski, A. V. Krasheninnikov, A. Tolvanen, K. Nordlund, and J. Keinonen, Physical Review B **81**, 153401 (2010).

[279] M. M. Ugeda, I. Brihuega, F. Guinea, and J. M. Gómez-Rodríguez, Physical Review Letters **104**, 096804 (2010).

[280] L. G. Cançado, A. Jorio, E. H. M. Ferreira, F. Stavale, C. A. Achete, R. B. Capaz, M. V. O. Moutinho, A. Lombardo, T. S. Kulmala, and A. C. Ferrari, Nano Letters **11**, 3190 (2011).

[281] A. Eckmann, A. Felten, A. Mishchenko, L. Britnell, R. Krupke, K. S. Novoselov, and C. Casiraghi, Nano Letters **12**, 3925 (2012).

[282] S. Plimpton, Journal of Computational Physics **117**, 1 (1995).

[283] D. W. Brenner, O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni, and S. B. Sinnott, Journal of Physics: Condensed Matter **14**, 783 (2002).

[284] R. Roldán, A. Fasolino, K. V. Zakharchenko, and M. I. Katsnelson, Physical Review B **83**, 174104 (2011).

[285] J. Haskins, A. Kınacı, C. Sevik, H. Sevinçli, G. Cuniberti, and T. Çağın, ACS Nano **5**, 3779 (2011).